All at Sea with Streams

Using Kafka to Detect Patterns in the Behaviour of Ships

#kafkasummit

@rmoff

Further reading

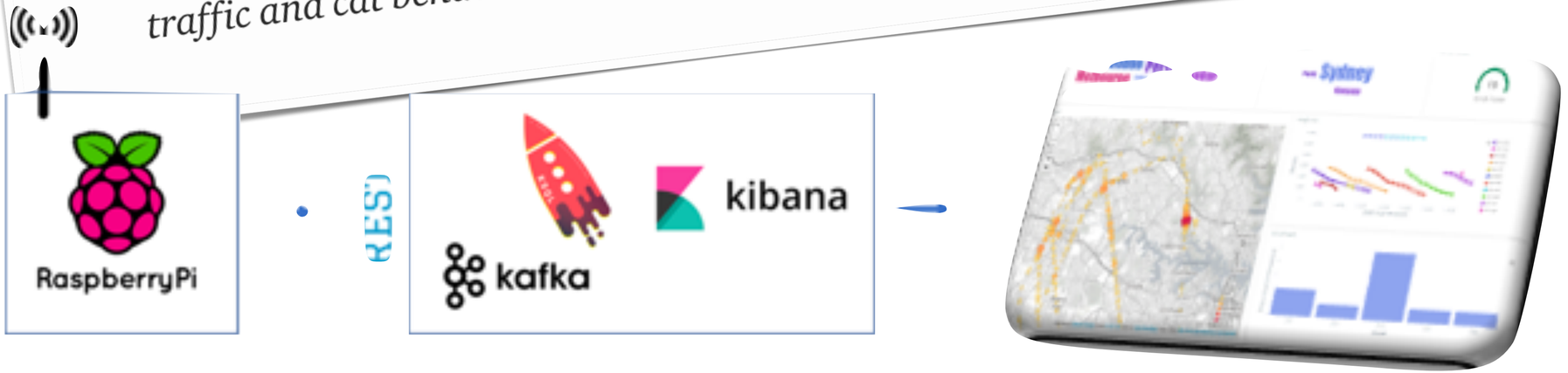
- This talk is built on concepts explored in detail in these two blogs:
  - https://www.confluent.co.uk/blog/streaming-etl-and-analytics-for-real-time-location-tracking/
  - https://www.confluent.co.uk/blog/streaming-data-with-confluent-and-ksqldb-for-new-use-cases-with-ais/

Planes

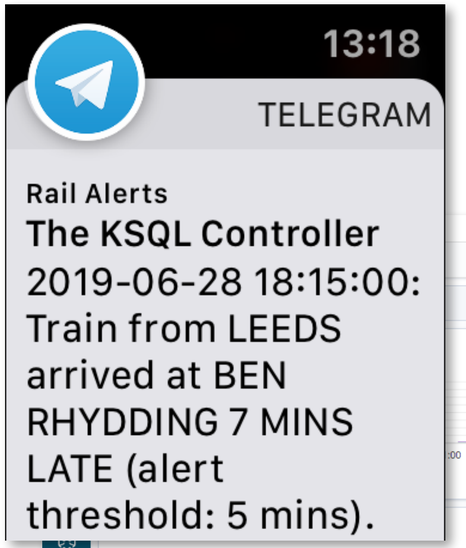
Using KSQL, Apache Kafka, a Raspberry Pi and a software defined radio to find the plane that wakes my cat

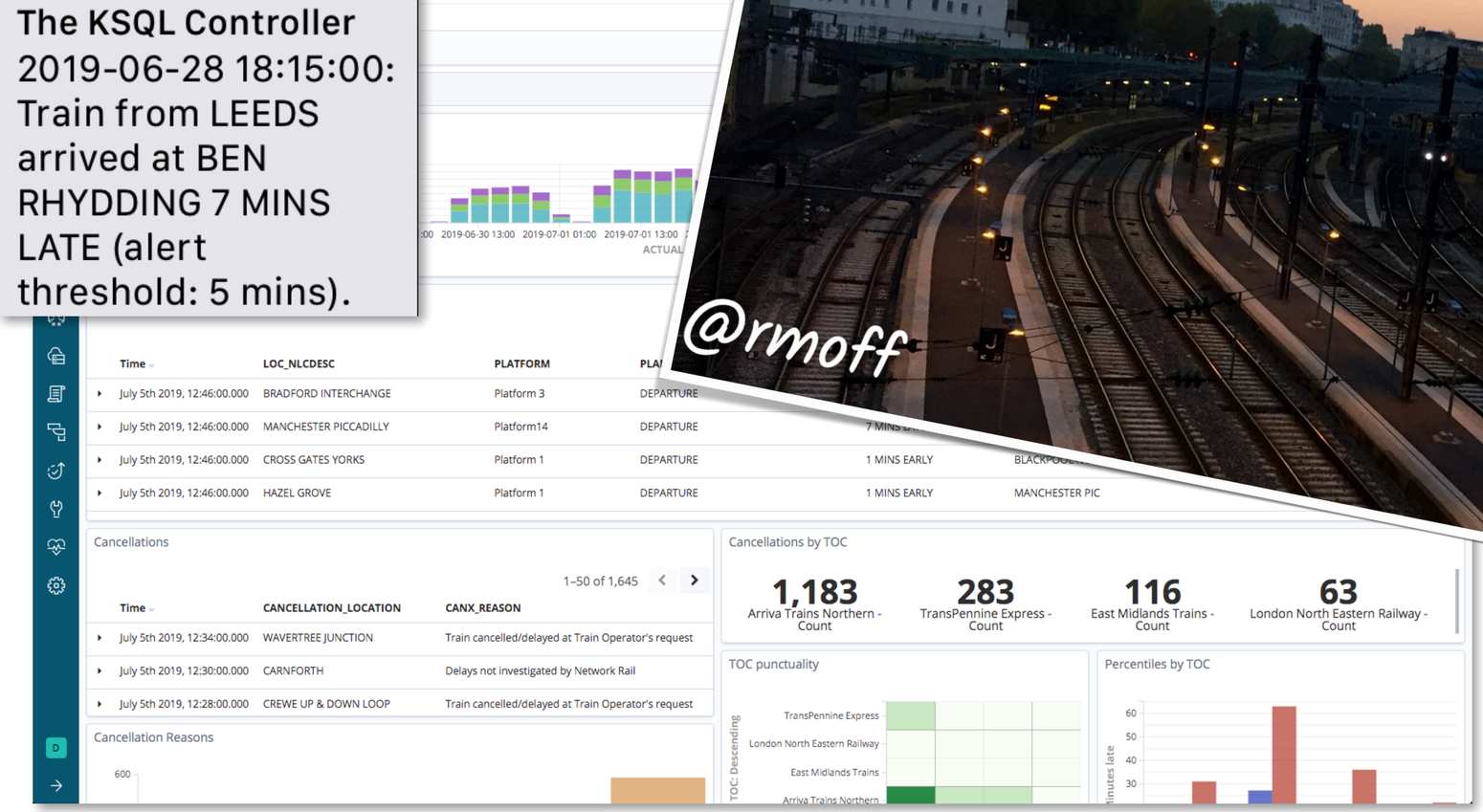
Simon Aubury May 27, 2018 · 5 min read

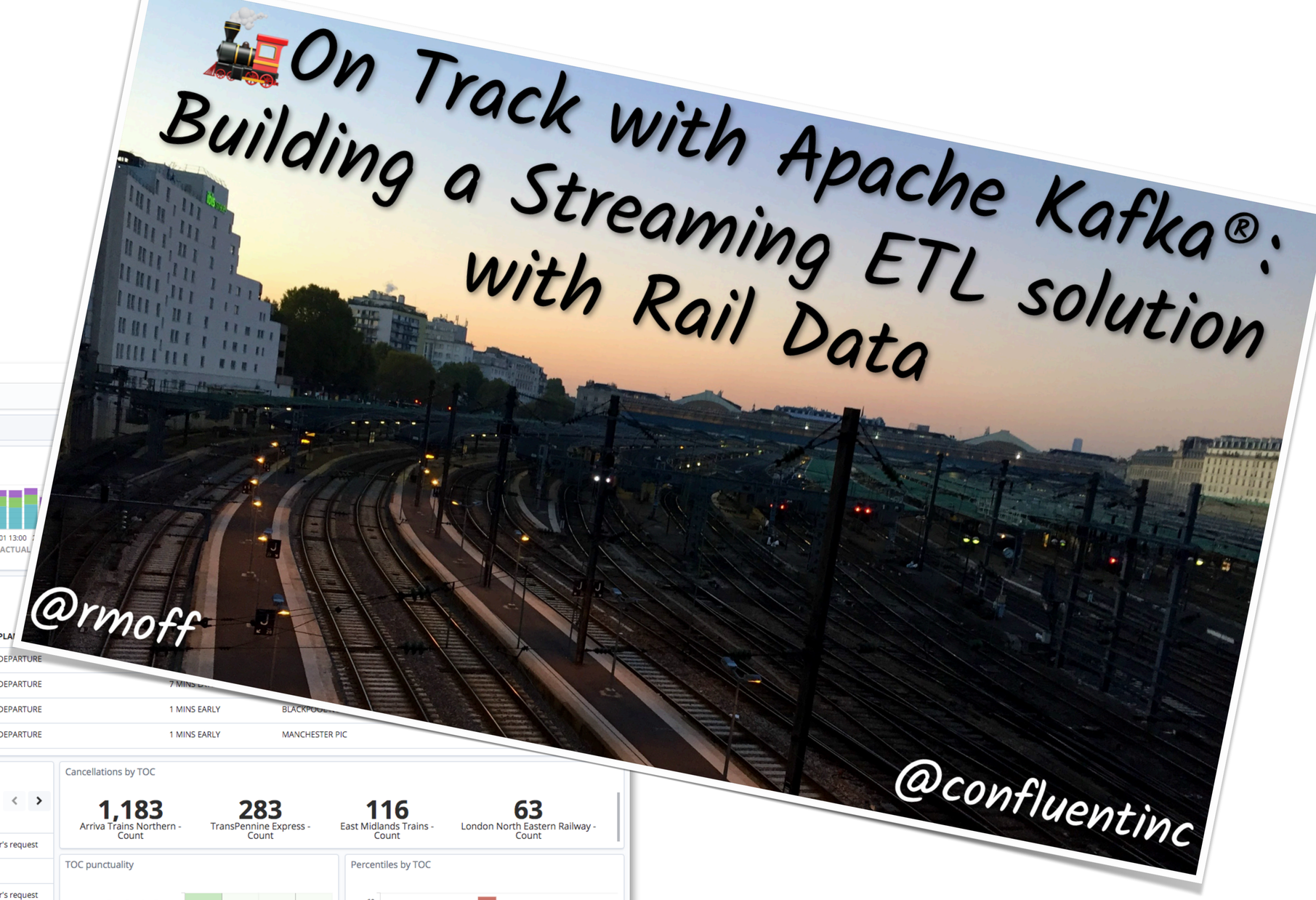
Using open-source streaming solutions to map the correlation between aircraft traffic and cat behaviour



Trains

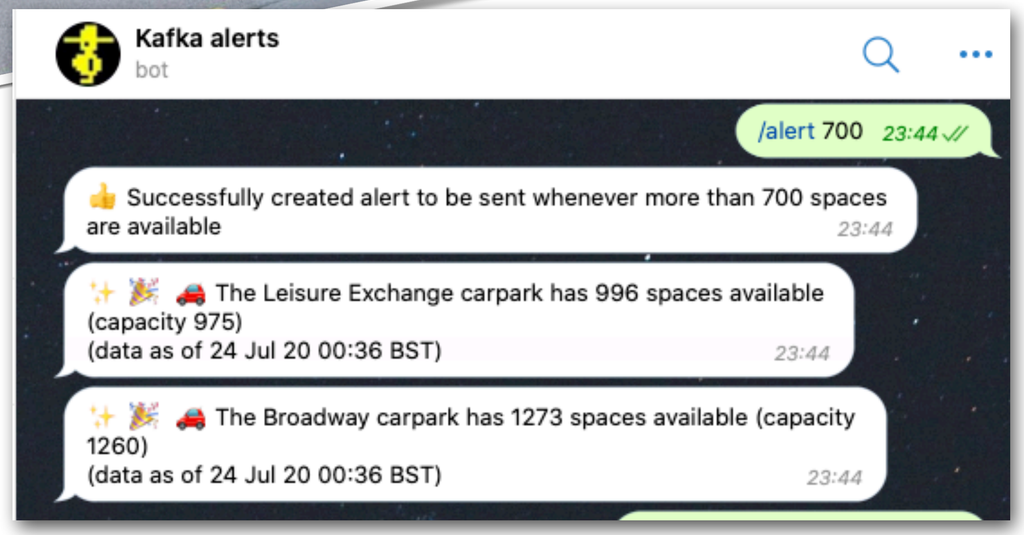

13:18
TELEGRAM
Rail Alerts
The KSQL Controller
2019-06-28 18:15:00: Train from LEEDS arrived at BEN RHYDDING 7 MINS LATE (alert threshold: 5 mins).


2019-06-30 13:00
2019-07-01 01:00
2019-07-01 13:00
ACTUAL
Time
LOC_NLCDESC
PLATFORM
PLA
July 5th 2019, 12:46:00.000
BRADFORD INTERCHANGE
Platform 3
DEPARTURE
July 5th 2019, 12:46:00.000
MANCHESTER PICCADILLY
Platform14
DEPARTURE
July 5th 2019, 12:46:00.000
CROSS GATES YORKS
Platform 1
DEPARTURE
1 MINS EARLY
July 5th 2019, 12:46:00.000
HAZEL GROVE
Platform 1
DEPARTURE
1 MINS EARLY
MANCHESTER PIC
Cancellations
1–50 of 1,645
Time
CANCELLATION_LOCATION
CANX_REASON
July 5th 2019, 12:34:00.000
WAVERTREE JUNCTION
Train cancelled/delayed at Train Operator's request
July 5th 2019, 12:30:00.000
CARNFORTH
Delays not investigated by Network Rail
July 5th 2019, 12:28:00.000
CREWE UP & DOWN LOOP
Train cancelled/delayed at Train Operator's request
Cancellation Reasons
600
Cancellations by TOC
1,183
Arriva Trains Northern - Count
283
TransPennine Express - Count
116
East Midlands Trains - Count
63
London North Eastern Railway - Count
TOC punctuality
TOC: Descending
TransPennine Express
London North Eastern Railway
East Midlands Trains
Arriva Trains Northern
Percentiles by TOC
Minutes late
60
50
40
30

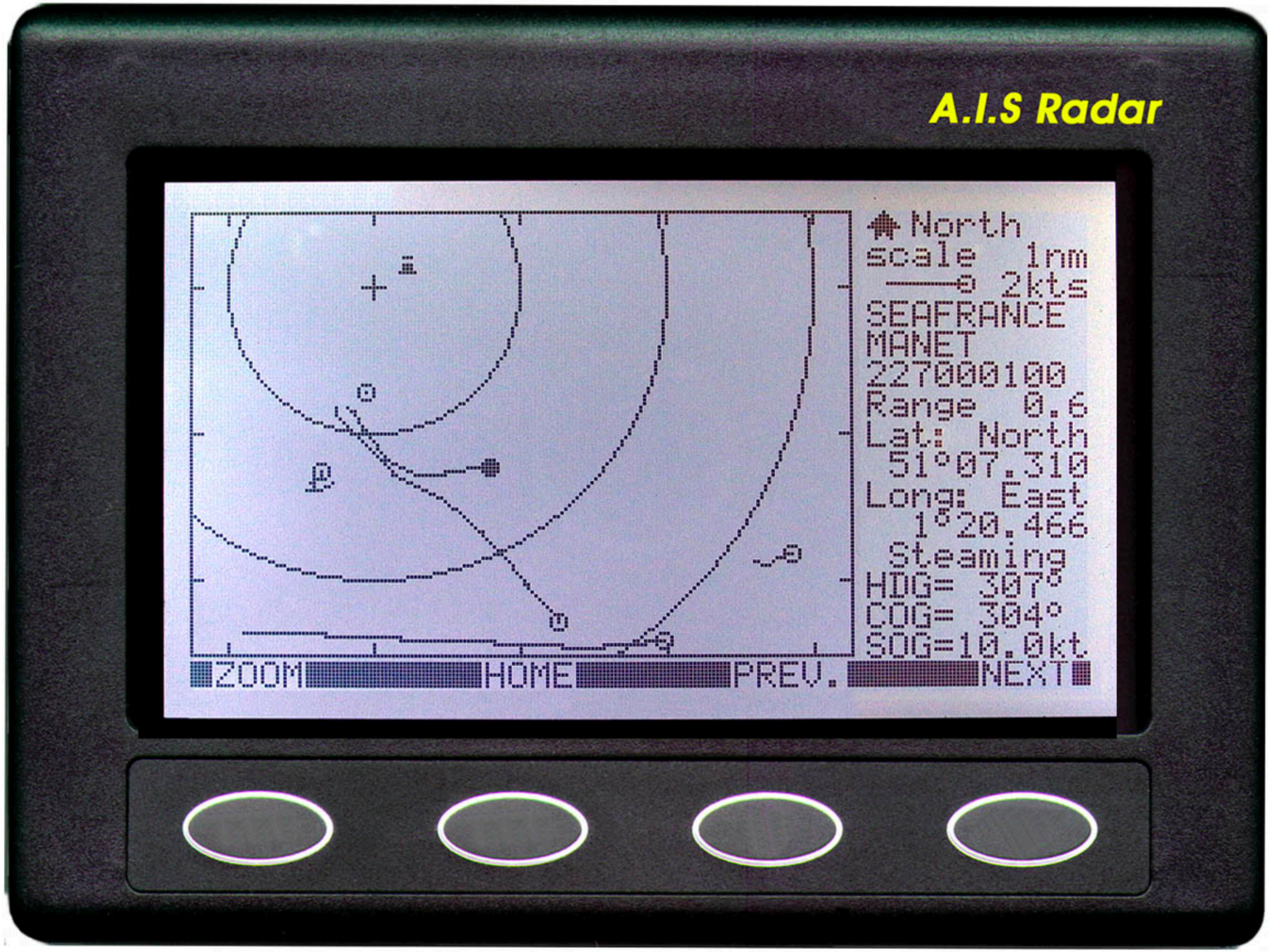

On Track with Apache Kafka®: Building a Streaming ETL solution with Rail Data
@rmoff
@confluentinc

Automobiles

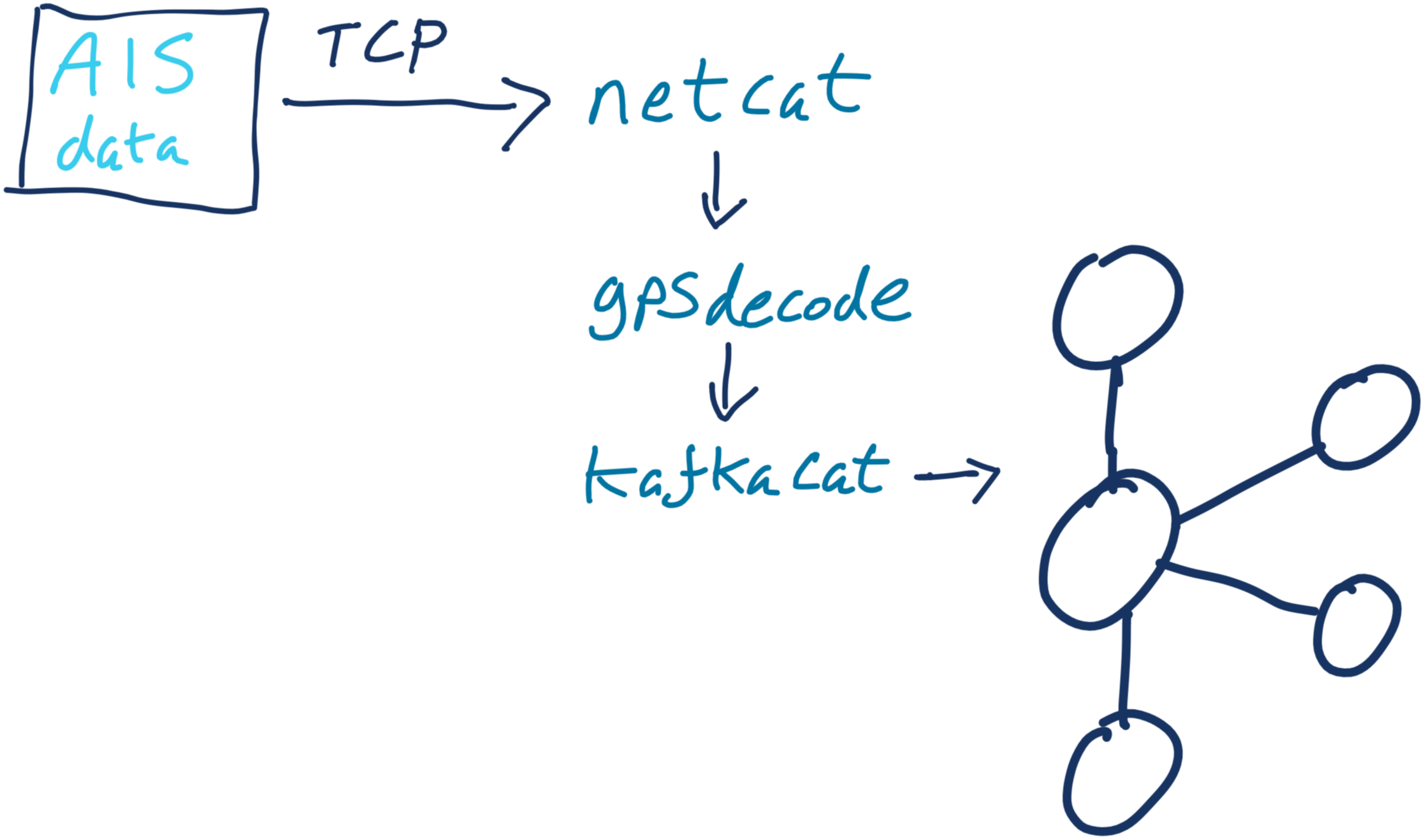
	capacity	emptyPlaces	pctFull
33Z	1132	921	18.6
17Z	611	474	22.4
4Z	98	63	35.7



...and Ships!


A.I.S Radar
North
scale 1nm
2kts
SEAFRANCE
MANET
227000100
Range 0.6
Lat: North
51°07.310
Long: East
1°20.466
Steaming
HDG= 307°
COG= 304°
SOG=10.0kt
ZOOM
HOME
PREV.
NEXT

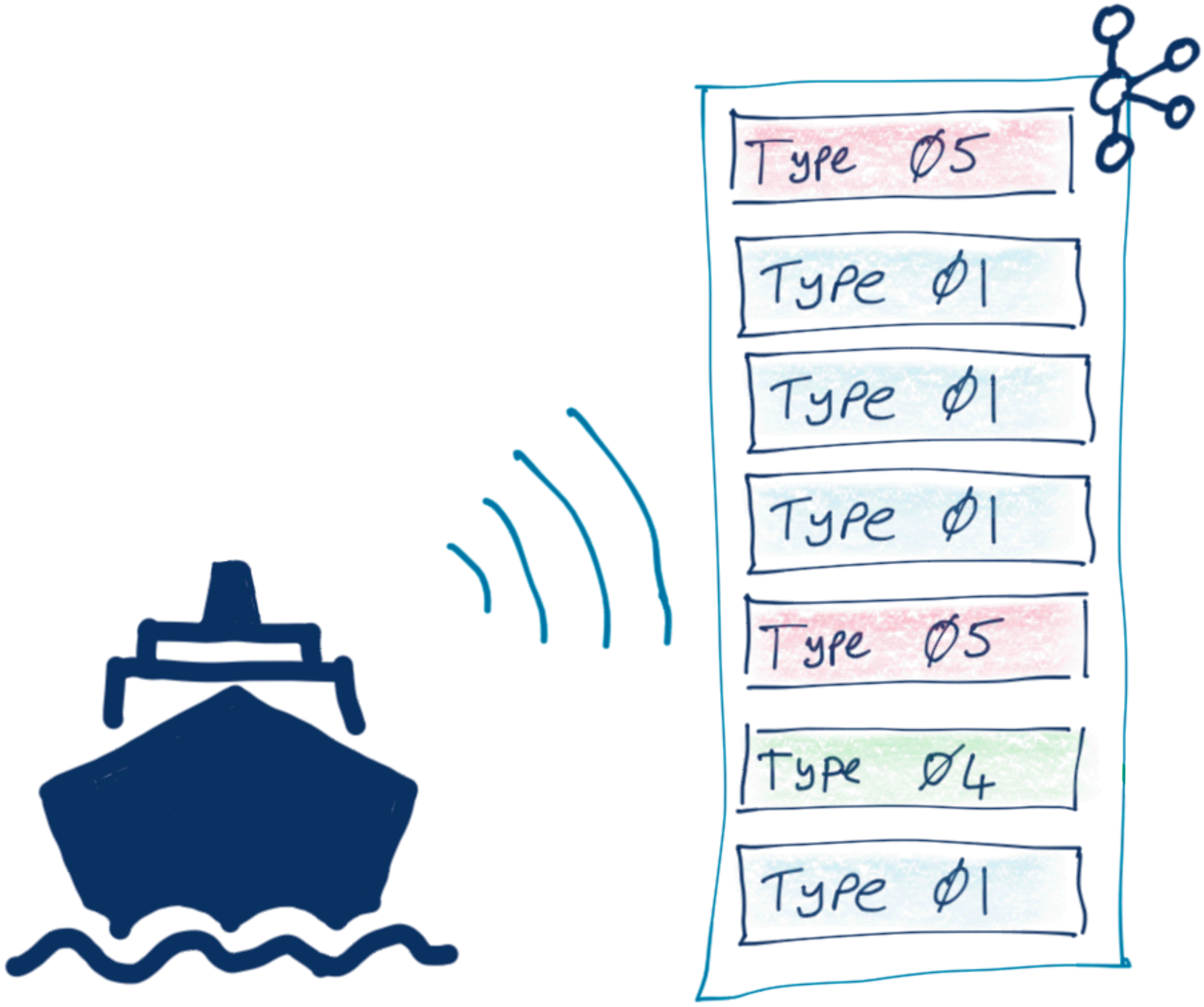
NASA MARINE
AIS ENGINE 3
Super sensitive 2 channel AIS receiver
with GPS Multiplexer
Receives all class A and class B messages
RoHS
COMPLIANT
VHF
ANTENNA
12 VOLT
INPUT
POWER
DATA
AIS DATA
OUTPUT

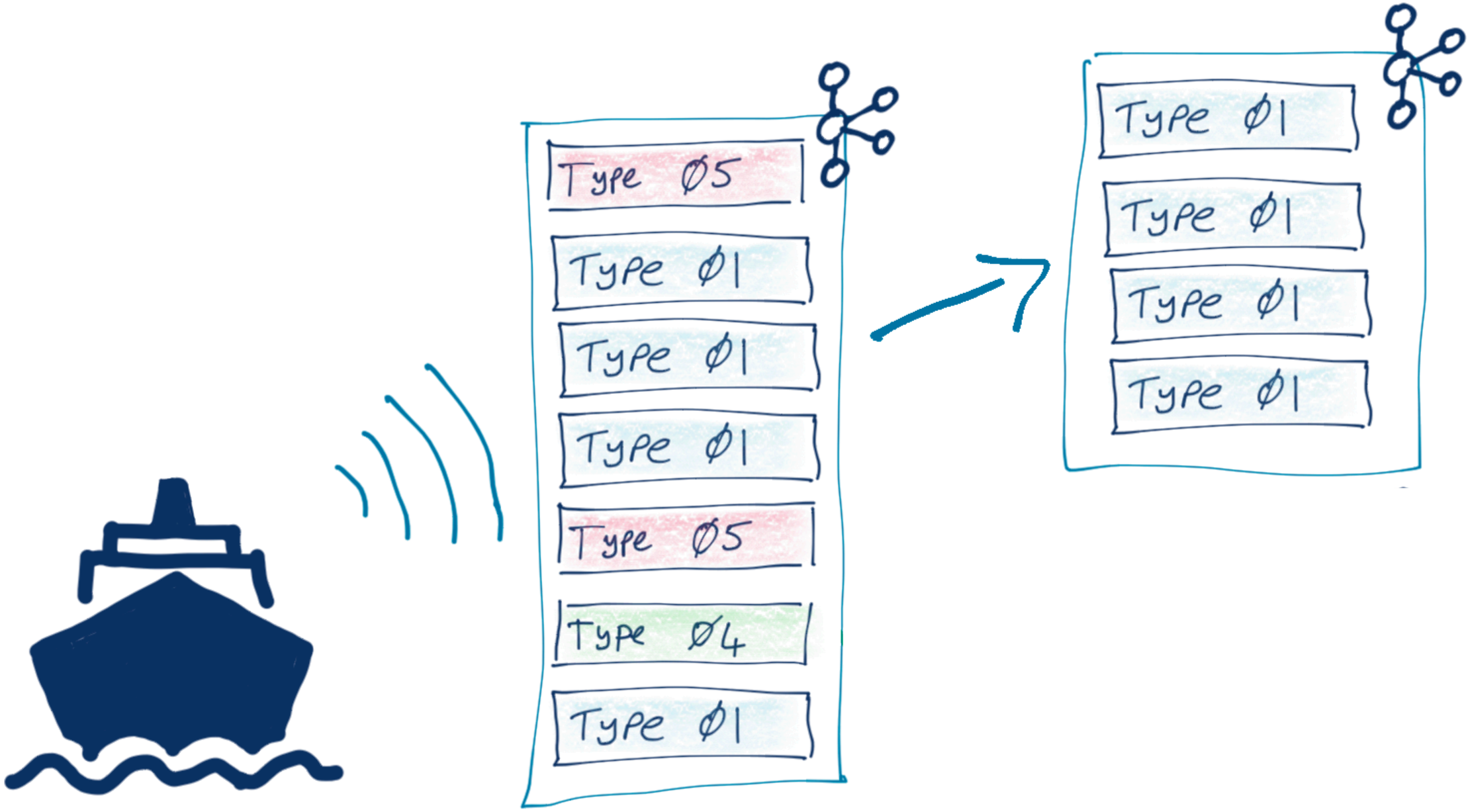

AIS data
TCP
netcat
gpsdecode
kafkacat

Handling different messages types in the same stream


```
{
  "class": "AIS",
  "device": "stdin",
  "type": 5,
  "repeat": 0,
  "mmsi": 219010873,
  "scaled": true,
  "imo": 0,
  "ais_version": 0,
  "callsign": "OXSH2",
  "shipname": "BELT SUPPLY",
  "shiptype": 70,
  "shiptype_text": "Cargo - all ships of this type",
  "to_bow": 6,
  "to_stern": 10,
  "to_port": 2,
  "to_starboard": 3,
  "epfd": 1,
  "epfd_text": "GPS",
  "eta": "11-24T00:51Z",
  "draught": 1.4,
  "destination": "SKAW ROUTE",
  "dte": 0
}
```

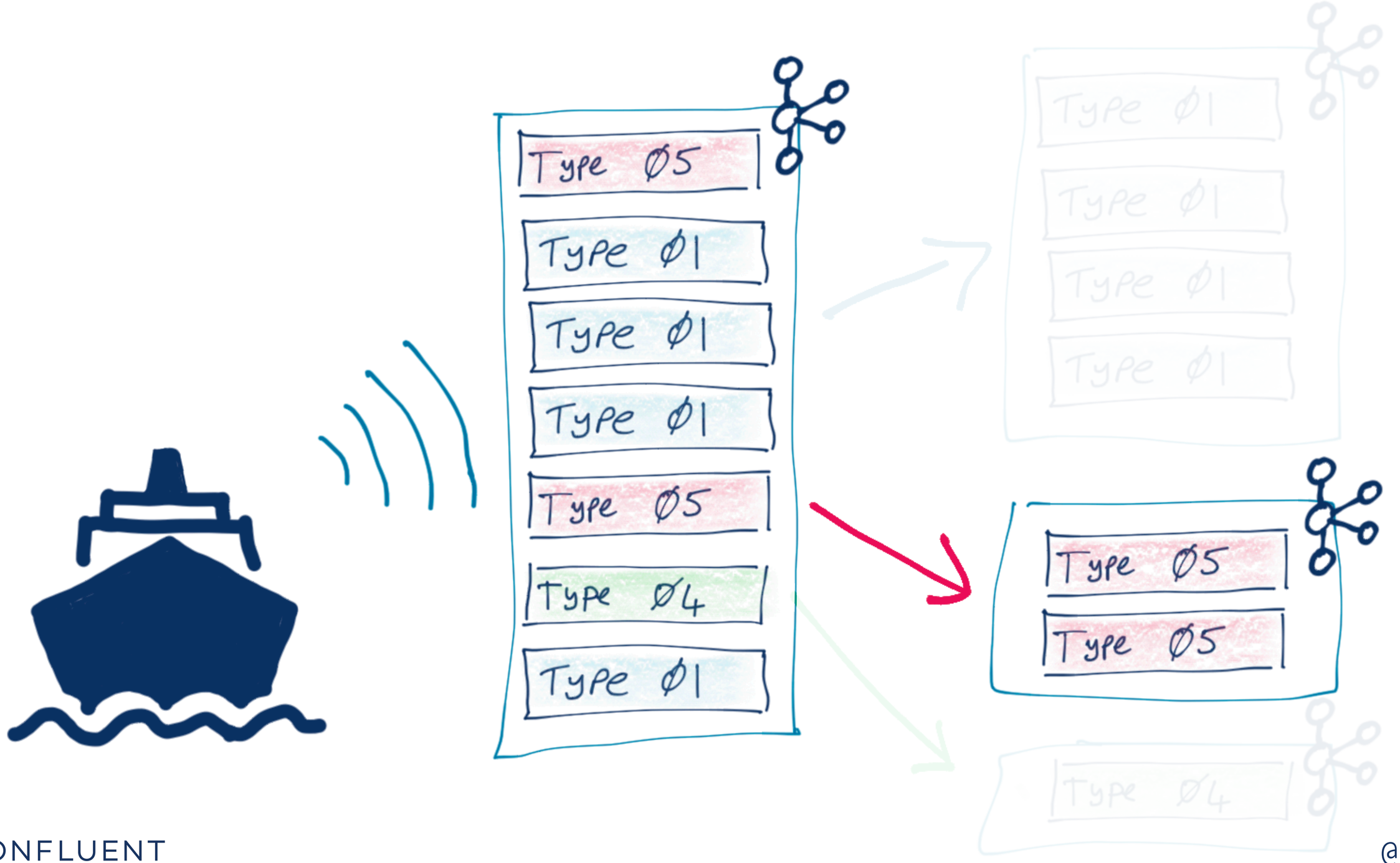
```
{
  "class": "AIS",
  "device": "stdin",
  "type": 1,
  "repeat": 0,
  "mmsi": 244675000,
  "scaled": true,
  "status": 0,
  "status_text": "Under way using engine",
  "turn": 0,
  "speed": 10.5,
  "accuracy": false,
  "lon": 4.004307,
  "lat": 59.000963,
  "course": 181.4,
  "heading": 180,
  "second": 10,
  "maneuver": 0,
  "raim": true,
  "radio": 33149
}
```


Type 05
Type 01
Type 01
Type 01
Type 05
Type 04
Type 01


Type Ø5
Type Ø1
Type Ø1
Type Ø1
Type Ø5
Type Ø4
Type Ø1
Type Ø1
Type Ø1
Type Ø1
Type Ø1

Use ksqlDB to route messages based on type

```
ksql> CREATE OR REPLACE STREAM AIS_MSG_TYPE_1_2_3 WITH (FORMAT='AVRO') AS
  SELECT  CAST(EXTRACTJSONFIELD(msg,'$.type') AS INT)             AS msg_type,
          CAST(EXTRACTJSONFIELD(msg,'$.mmsi') AS VARCHAR)         AS mmsi,
          CAST(EXTRACTJSONFIELD(msg,'$.status_text') AS VARCHAR)  AS status_text,
          CAST(EXTRACTJSONFIELD(msg,'$.speed') AS DOUBLE)         AS speed,
          CAST(EXTRACTJSONFIELD(msg,'$.course') AS DOUBLE)        AS course,
          CAST(EXTRACTJSONFIELD(msg,'$.heading') AS INT)          AS heading
  FROM    AIS_RAW
  WHERE   EXTRACTJSONFIELD(msg,'$.type') IN ('1' ,'2' ,'3' ,'18' ,'27')
  PARTITION BY CAST(EXTRACTJSONFIELD(msg,'$.mmsi') AS VARCHAR);
```

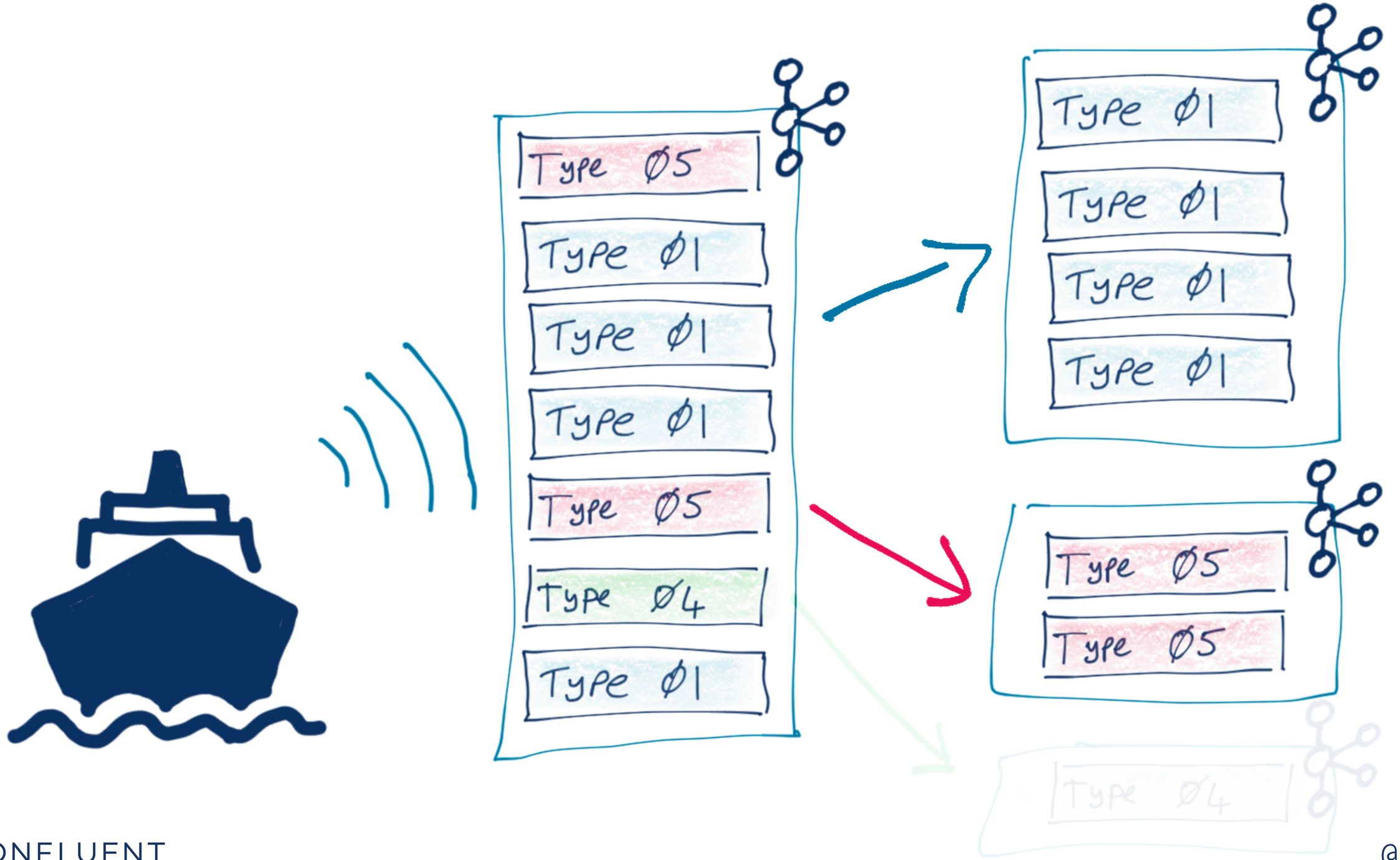

Type Ø5
Type Ø1
Type Ø1
Type Ø1
Type Ø5
Type Ø4
Type Ø1
Type Ø1
Type Ø1
Type Ø1
Type Ø1
Type Ø5
Type Ø5
Type Ø4

```
ksql> CREATE OR REPLACE STREAM AIS_MSG_TYPE_5 WITH (FORMAT='AVRO')          AS
  SELECT  CAST(EXTRACTJSONFIELD(msg,'$.type') AS INT)                   AS msg_type,
          CAST(EXTRACTJSONFIELD(msg,'$.mmsi') AS VARCHAR)               AS mmsi,
          CAST(EXTRACTJSONFIELD(msg,'$.callsign') AS VARCHAR)           AS callsign,
          CAST(EXTRACTJSONFIELD(msg,'$.shipname') AS VARCHAR)           AS shipname_raw,
          CONCAT(CAST(EXTRACTJSONFIELD(msg,'$.shipname') AS VARCHAR),
                 ' (',
                 CAST(EXTRACTJSONFIELD(msg,'$.callsign') AS VARCHAR),
                 ')')                                                   AS shipname,
          CAST(EXTRACTJSONFIELD(msg,'$.shiptype_text') AS VARCHAR)      AS shiptype_text,
          CAST(EXTRACTJSONFIELD(msg,'$.destination') AS VARCHAR)        AS destination
  FROM    AIS_RAW
  WHERE   EXTRACTJSONFIELD(msg,'$.type') = '5'
  PARTITION BY CAST(EXTRACTJSONFIELD(msg,'$.mmsi') AS VARCHAR);
```

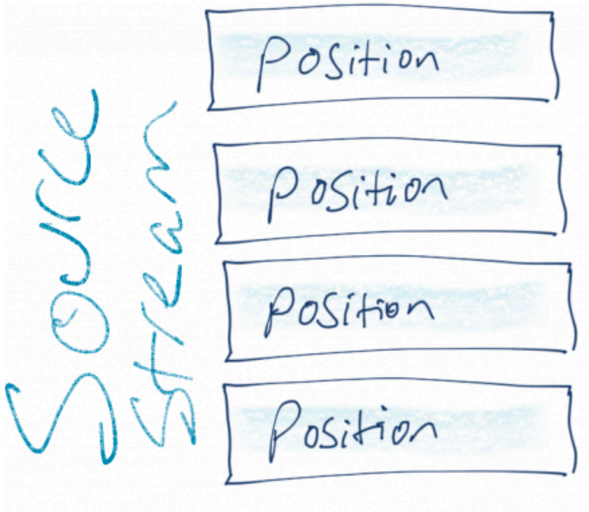
Build a materialized view of state from a Kafka topic

```
ksql> CREATE TABLE SHIP_INFO AS
  SELECT MMSI,
        MAX(ROWTIME) AS LAST_INFO_PING_TS,
        LATEST_BY_OFFSET(SHIPNAME) AS SHIPNAME,
        LATEST_BY_OFFSET(DRAUGHT) AS DRAUGHT,
        LATEST_BY_OFFSET(DESTINATION) AS DESTINATION
    FROM AIS_MSG_TYPE_5
  GROUP BY MMSI
  EMIT CHANGES;

ksql> SELECT MMSI,
       TIMESTAMPTOSTRING(LAST_INFO_PING_TS,'HH:mm:ss','Europe/London') AS LAST_INFO_PING_TS,
       SHIPNAME,
       DRAUGHT,
       DESTINATION
  FROM SHIP_INFO
  WHERE MMSI = '255805587';
+----------+------------------+------------------+--------+------------+
|MMSI      |LAST_INFO_PING_TS |SHIPNAME          |DRAUGHT |DESTINATION |
+----------+------------------+------------------+--------+------------+
|255805587 |11:17:17          |NCL AVEROY (CQHL) |7.5     |SVELGEN     |
+----------+------------------+------------------+--------+------------+
```


Type Ø5
Type Ø1
Type Ø1
Type Ø1
Type Ø5
Type Ø4
Type Ø1
Type Ø1
Type Ø1
Type Ø1
Type Ø1
Type Ø5
Type Ø5
Type Ø4

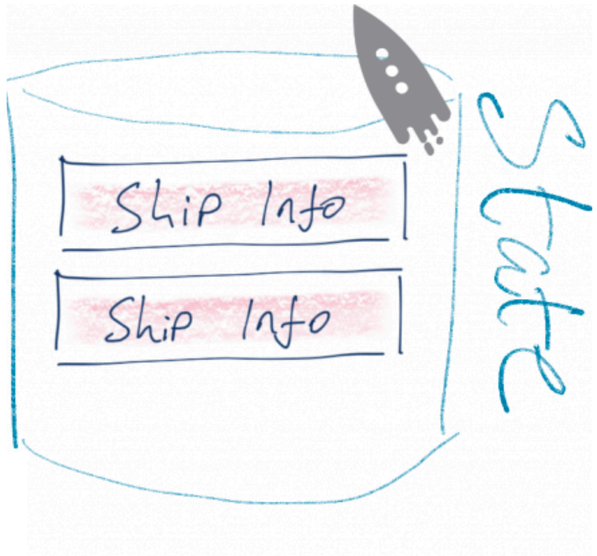
AIS_MSG_TYPE_1_2_3

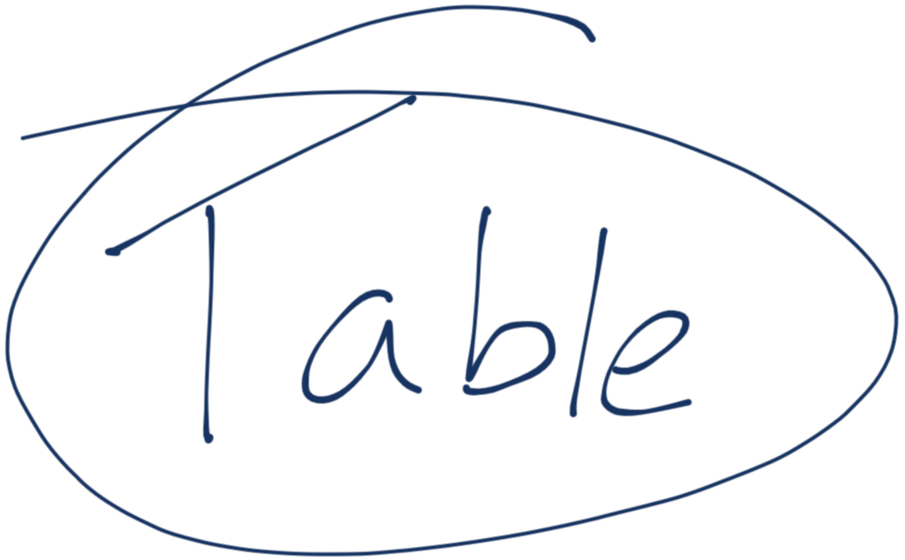


Stream

STATUS_TS	SHIP_INFO_A_MMSI	STATUS_REPORT_HEAD…	STATUS_REPORT_LAT	STATUS_REPORT_LON	STATUS_REPORT_SPEE…	STATUS_REPORT_…
Jul 20, 2021 @ 09:…	257046470	247	68.089	13.221	0	Engaged in fis…
Jul 20, 2021 @ 09:…	232015045	179	58.439	5.99	0	Under way usin…
Jul 20, 2021 @ 09:…	258992480	191	59.214	10.536	0	Under way usin…
Jul 20, 2021 @ 09:…	257399400	272	65.81	12.47	10.8	Under way usin…
Jul 20, 2021 @ 09:…	257350400	220	70.032	20.846	0	Under way usin…
Jul 20, 2021 @ 09:…	257107040	245	63.049	7.424	0	Engaged in fis…
Jul 20, 2021 @ 09:…	258051000	511	69.705	19.026	0	Under way usin…
Jul 20, 2021 @ 09:…	257063250	228	64.388	10.514	0.2	Under way sail…
Jul 20, 2021 @ 09:…	258534000	66	61.106	5.729	0.1	Under way usin…
Jul 20, 2021 @ 09:…	259321000	229	68.154	15.37	9	Under way usin…
Jul 20, 2021 @ 09:…	636016273	90	57.793	9.502	11.7	Under way usin…
Jul 20, 2021 @ 09:…	538002425	44	70.968	18.958	12.5	Under way usin…
Jul 20, 2021 @ 09:…	259620000	131	61.355	4.889	0	Under way usin…
Jul 20, 2021 @ 09:…	565798000	331	60.444	5.024	4.8	Moored
Jul 20, 2021 @ 09:…	257127460	102	65.775	12.247	0.1	Under way usin…
Jul 20, 2021 @ 09:…	257298000	64	66.022	12.628	0	Under way usin…
Jul 20, 2021 @ 09:…	219023738	343	58.544	10.506	2.1	Not defined
Jul 20, 2021 @ 09:…	257013420	511	68.256	14.267	0	Engaged in fis…

SHIP_INFO

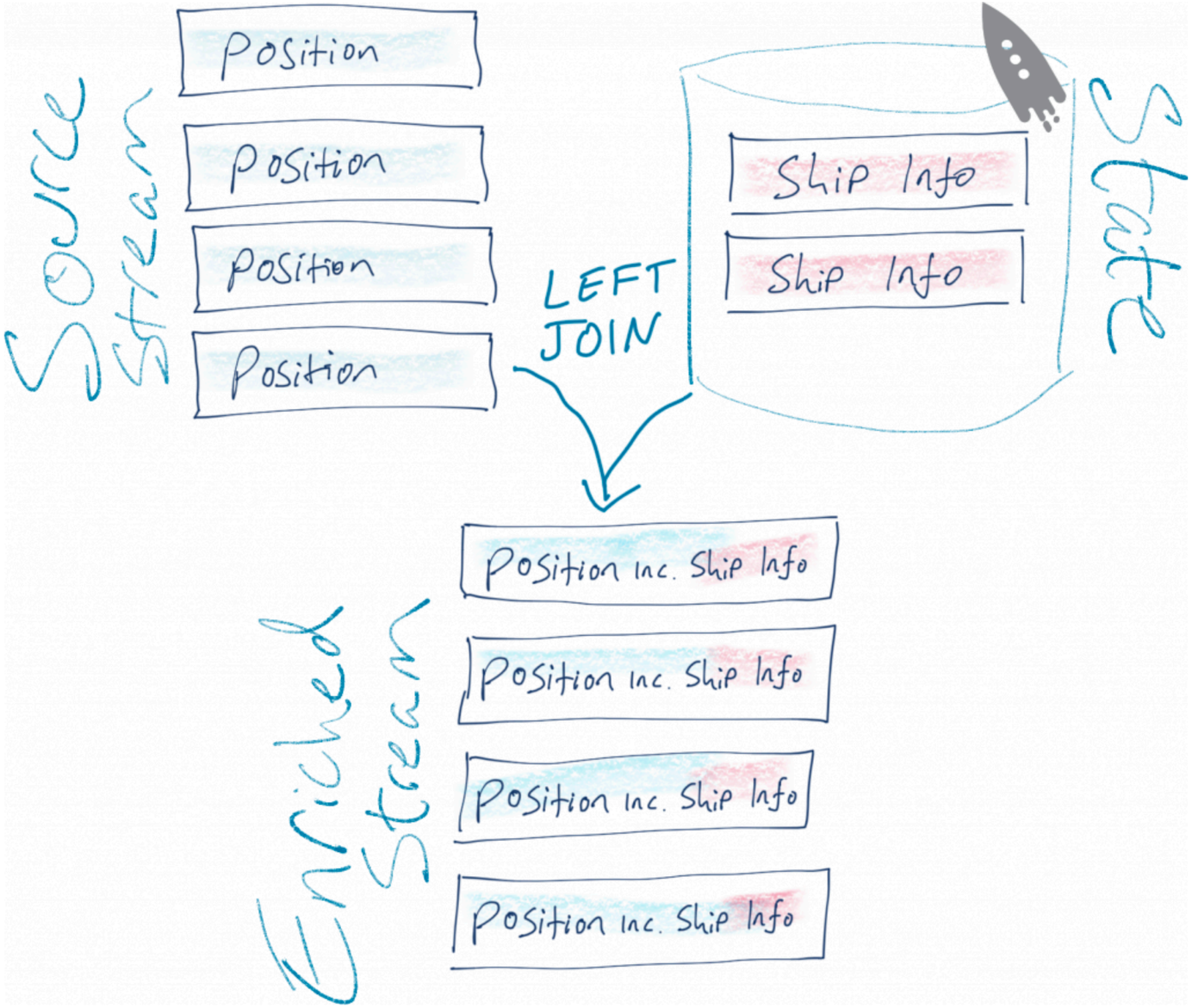





```
ksql> SELECT SHIPNAME, SHIPTYPE_TEXT, SHIP_AREA_M_SQ, DRAUGHT, DESTINATION_LIST
>       FROM SHIP_INFO
>      WHERE MMSI IN ('304604000','377134000','229631000');
+----------------------------+----------------------------+----------------------------+
|SHIPNAME                    |SHIPTYPE_TEXT               |SHIP_AREA_M_SQ              |
+----------------------------+----------------------------+----------------------------+
|LAKE ERIE (V2CF1)           |Cargo - No additional informatio|4440                    |
|                            |n                           |                            |
|SELVAAGSUND (J8B3801)       |Cargo - all ships of this type  |902                     |
|CIELO DI NEW YORK (9HA3466) |Tanker - all ships of this type |5856                    |
Query terminated
ksql>
```

AIS_MSG_TYPE_1_2_3

SHIP_INFO



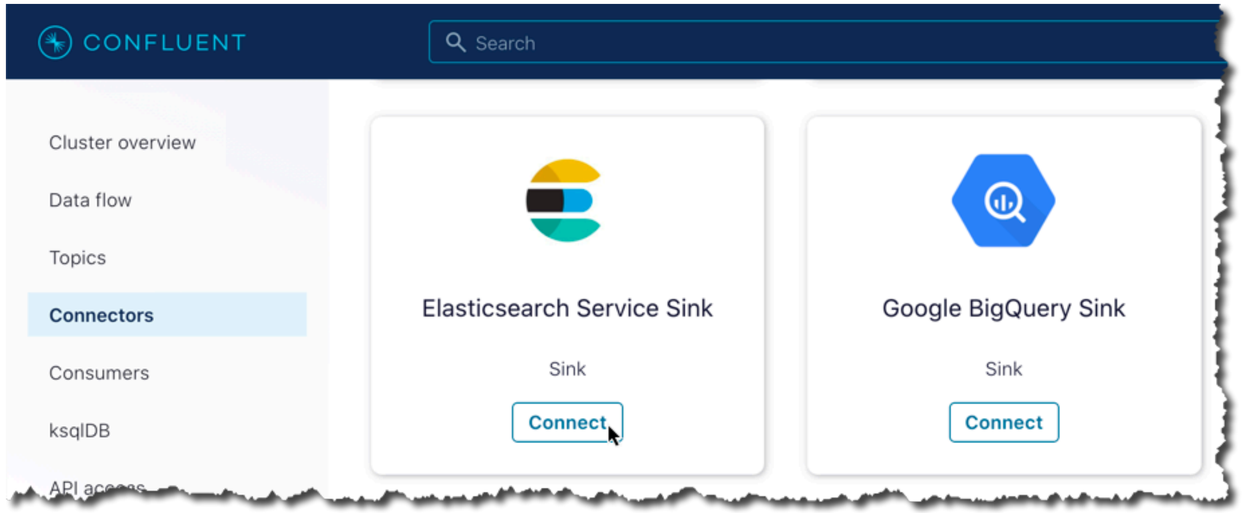
Join the stream to the table

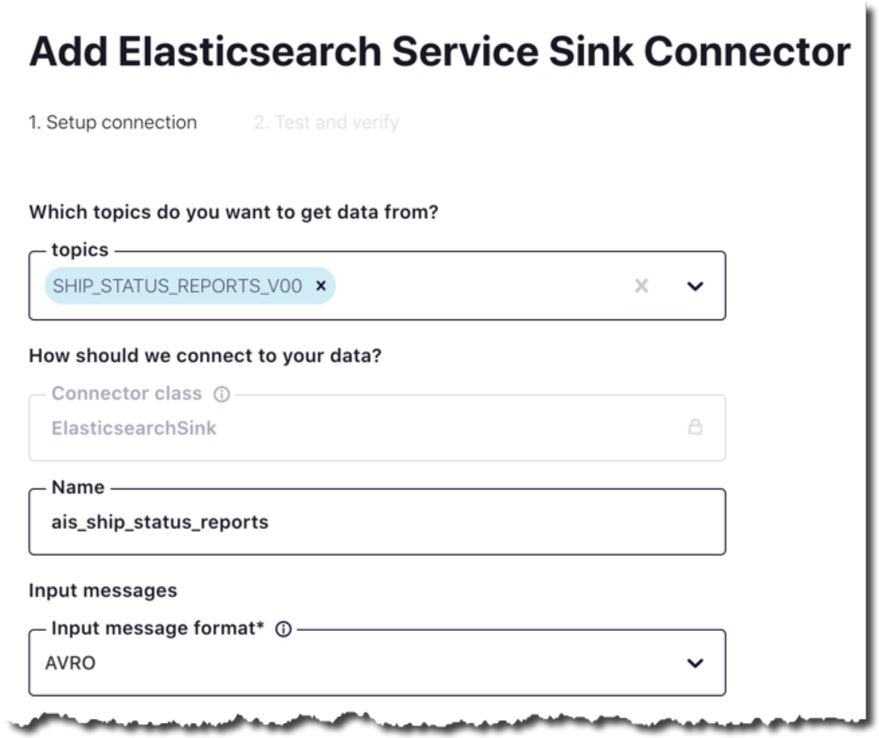
```
ksql> CREATE STREAM SHIP_STATUS_REPORTS WITH
  (KAFKA_TOPIC='SHIP_STATUS_REPORTS_V00') AS
SELECT STATUS_REPORT.ROWTIME AS STATUS_TS,
       STATUS_REPORT.*,
       SHIP_INFO.*
FROM  AIS_MSG_TYPE_1_2_3 STATUS_REPORT
      LEFT JOIN SHIP_INFO SHIP_INFO
        ON STATUS_REPORT.MMSI = SHIP_INFO.MMSI
;
```

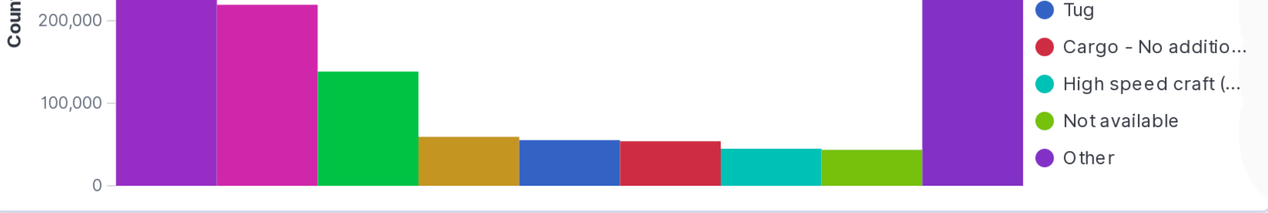
Ship movements (msg type 1) enriched with ship details (msg type 5)

```
ksql> SELECT TIMESTAMPTOSTRING(STATUS_TS,'HH:mm:ss','Europe/Oslo') AS STATUS_TS,
             SHIP_LOCATION,
             STATUS_REPORT_STATUS_TEXT,
             SHIP_INFO_SHIPNAME,
             SHIP_INFO_DRAUGHT,
             SHIP_INFO_DESTINATION_LIST
        FROM SHIP_STATUS_REPORTS
       WHERE SHIP_INFO_MMSI = '255805587'
       EMIT CHANGES;
+----------+------------------------------+--------------------------+--------------------+------------------+---------------------+
|STATUS_TS |SHIP_LOCATION                 |STATUS_REPORT_STATUS_TEXT |SHIP_INFO_SHIPNAME  |SHIP_INFO_DRAUGHT |SHIP_INFO_DESTINATION|
+----------+------------------------------+--------------------------+--------------------+------------------+---------------------+
|11:37:47  |{lat=61.773223, lon=5.294023} |Moored                    |NCL AVEROY (CQHL)   |7.5               |[SVELGEN]            |
[…]
|17:16:45  |{lat=61.939807, lon=5.143242} |Under way using engine    |NCL AVEROY (CQHL)   |7.5               |[FLORO]              |
[…]
|23:05:25  |{lat=62.468148, lon=6.137387} |Under way using engine    |NCL AVEROY (CQHL)   |8.1               |[ALESUND]            |
[…]
|23:11:04  |{lat=62.468122, lon=6.13745}  |Under way using engine    |NCL AVEROY (CQHL)   |8.1               |[ALESUND]            |
[…]
|23:35:47  |{lat=62.468125, lon=6.137473} |Moored                    |NCL AVEROY (CQHL)   |8.1               |[ALESUND]            |
[…]
```

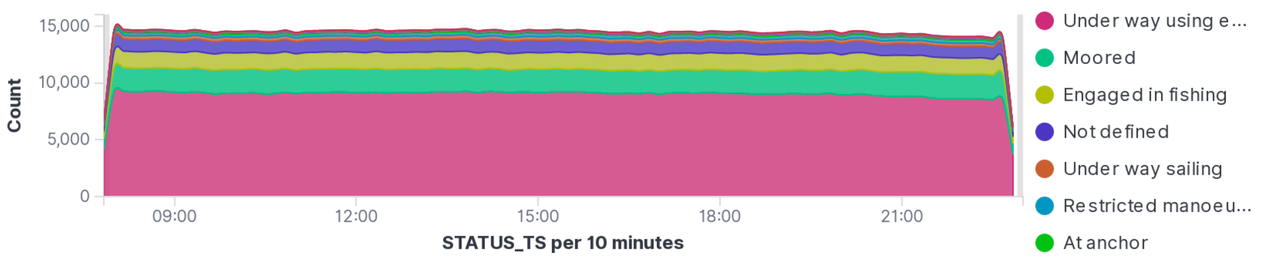
Stream the enriched Kafka topic to Elasticsearch



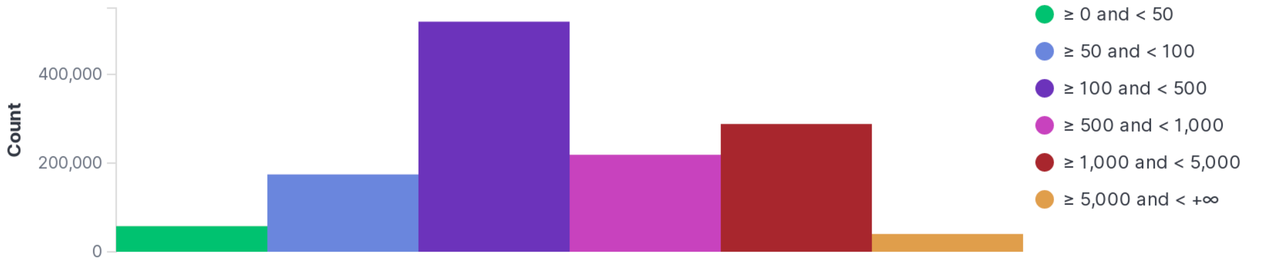


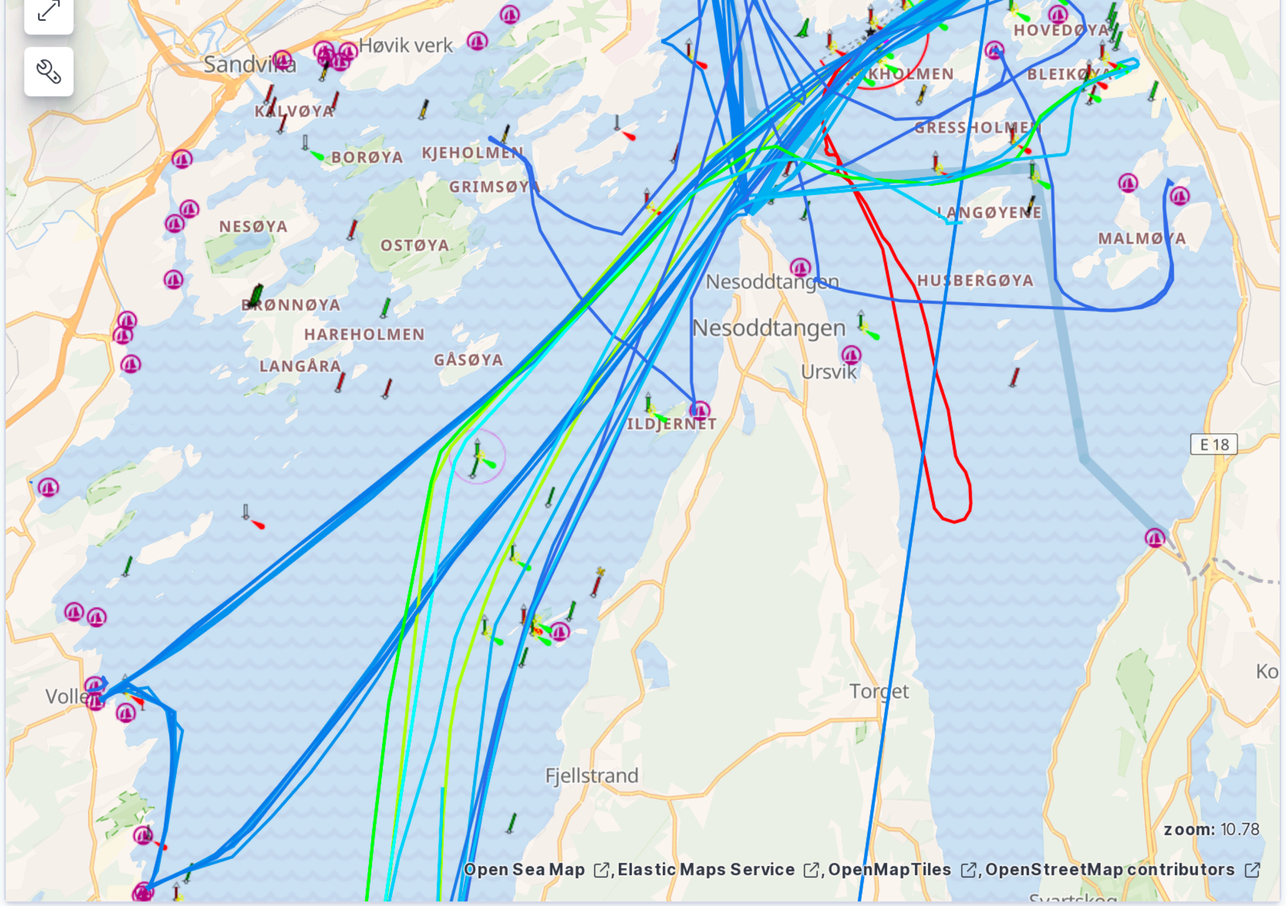


Status over time



Ship Area m²



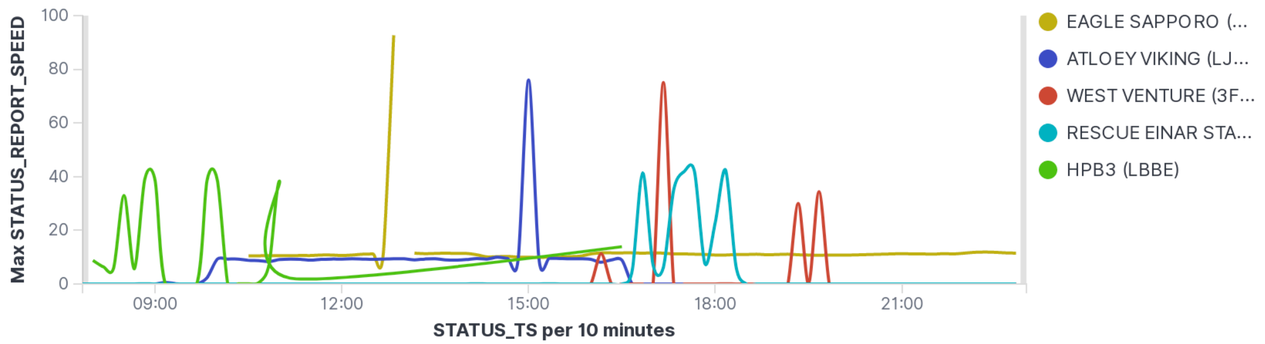


Status Reports

1–50 of 1333809

Time	SHIP_INFO_SHIPNAME	SHIP_INFO_DESTINATION	SHIP_INFO_SHIPTYPE_TEXT
Mar 4, 2021 @ 22:54:19.658	-	-	-
Mar 4, 2021 @ 22:54:19.658	BUKKEN (LENN)	LEPSOEYREVET	Not available
Mar 4, 2021 @ 22:54:19.658	BIOSTAR (LDBB)	FISHFARMS	Cargo - all ships of this type
Mar 4, 2021 @ 22:54:19.658	MULTI PROVIDER (LEYQ)	CH16	Tug
Mar 4, 2021 @ 22:54:19.658	SVANFJELL ()		Other Type - no additional informati

Top five fastest ships



AIS shippping data published the Norwegian Coastal Administration under Norwegian licence for Open Government data (NLOD)

Identifying Transshipping Behaviour

frontiers in Marine Science

ORIGINAL RESEARCH
published: 23 July 2018
doi: 10.3389/fmars.2018.00240

Identifying Global Patterns of Transshipment Behavior

Nathan A. Miller[1], Aaron Roan[2], Timothy Hochberg[3], John Amos[1] and David A. Kroodsma[3]*

[1] SkyTruth, Shepherdstown, WV, United States, [2] Google, Mountain View, CA, United States, [3] Global Fishing Watch, Washington, DC, United States

Transshipment at sea, the offloading of catch from a fishing vessel to a refrigerated vessel far from port, can obscure the actual source of the catch, complicating sustainable fisheries management, and may allow illegally caught fish to enter the legitimate seafood market. Transshipment activities often occur in regions of unclear jurisdiction where policymakers or enforcement agencies may be slow to act against a challenge they cannot see. To address this limitation, we processed 32 billion Automatic Identification System (AIS) messages from ocean-going vessels from 2012 to the end of 2017 and identified and tracked 694 cargo vessels capable of transshipping at sea and transporting fish (referred to as transshipmen
vessels loitered at sea long enc
where we see a fishing vessel
engage in transshipment. We f
and flag states exhibiting limited
seas and 42% involve vessels fl
high seas is relatively common,
seas having at least one encour
analysis reveals that addressing
trafficking, bonded labor) assoc
perspective and transnational c

Keywords: fisheries, transshipment, refri

OPEN ACCESS

Edited by:
Dyhia Belhabib,
University of British Columbia, Canada

Reviewed by:
Brett W. Molony,
Department of Primary Industries and Regional Development of Western Australia (DPIRD), Australia
Simon Foale,
James Cook University, Australia

****Correspondence:***
Nathan A. Miller
nate@skytruth.org

Specialty section:
This article was submitted to Marine Fisheries, Aquaculture and Living Resources, a section of the journal Frontiers in Marine Science

Received: *31 March 2018*
Accepted: *25 June 2018*
Published: *23 July 2018*

Citation:
Miller NA, Roan A, Hochberg T, Amos J and Kroodsma DA (2018) Identifying Global Patterns of Transshipment Behavior. Front. Mar. Sci. 5:240. doi: 10.3389/fmars.2018.00240

INTRODUCTION

A transshipment occurs when two
between vessels at sea and far fro
sea and remain on the fishing grou
moves product to market more effi
traceability and transparency in th
(FAO) of the United Nations estim
unregulated (IUU) (Food Agricultu
has also been linked to human tr
indefinitely, resulting in de facto slav
the industry can access reputable ma
transshipments (Environmental Ju
and prevents an accurate measurem
limiting our ability to fish the ocean

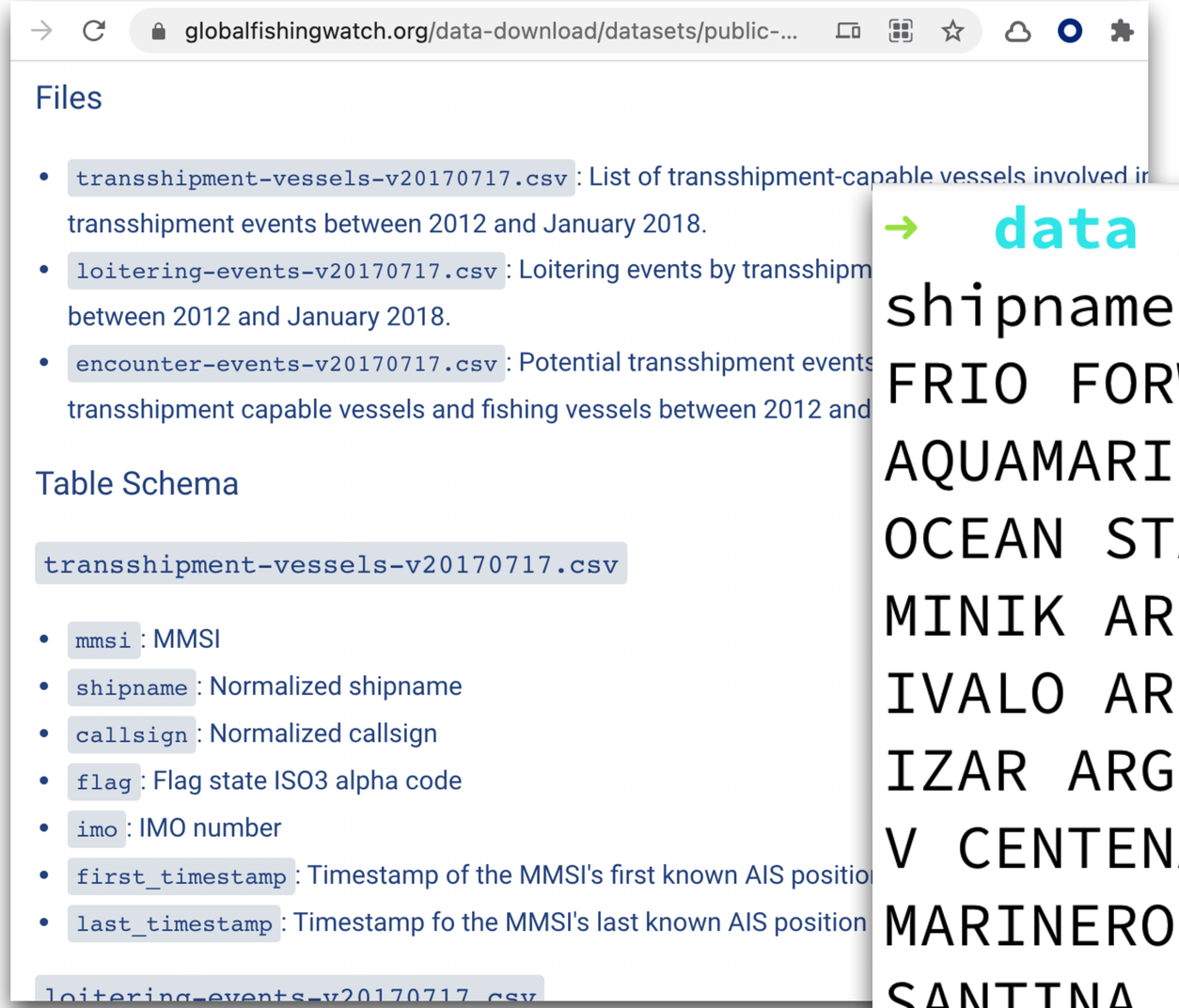
Two-Vessel Encounter

Encounters were identified from AIS data as locations where a fishing vessel and a transshipment vessel were continuously within 500 m for at least 2 h and traveling at <2 knots, while at least 10 km from a coastal anchorage.

Miller NA, Roan A, Hochberg T, Amos J and Kroodsma DA (2018) Identifying Global Patterns of Transshipment Behavior. Front. Mar. Sci. 5:240. doi: 10.3389/fmars.2018.00240

Credit: Francisco Blaha

Additional ship data



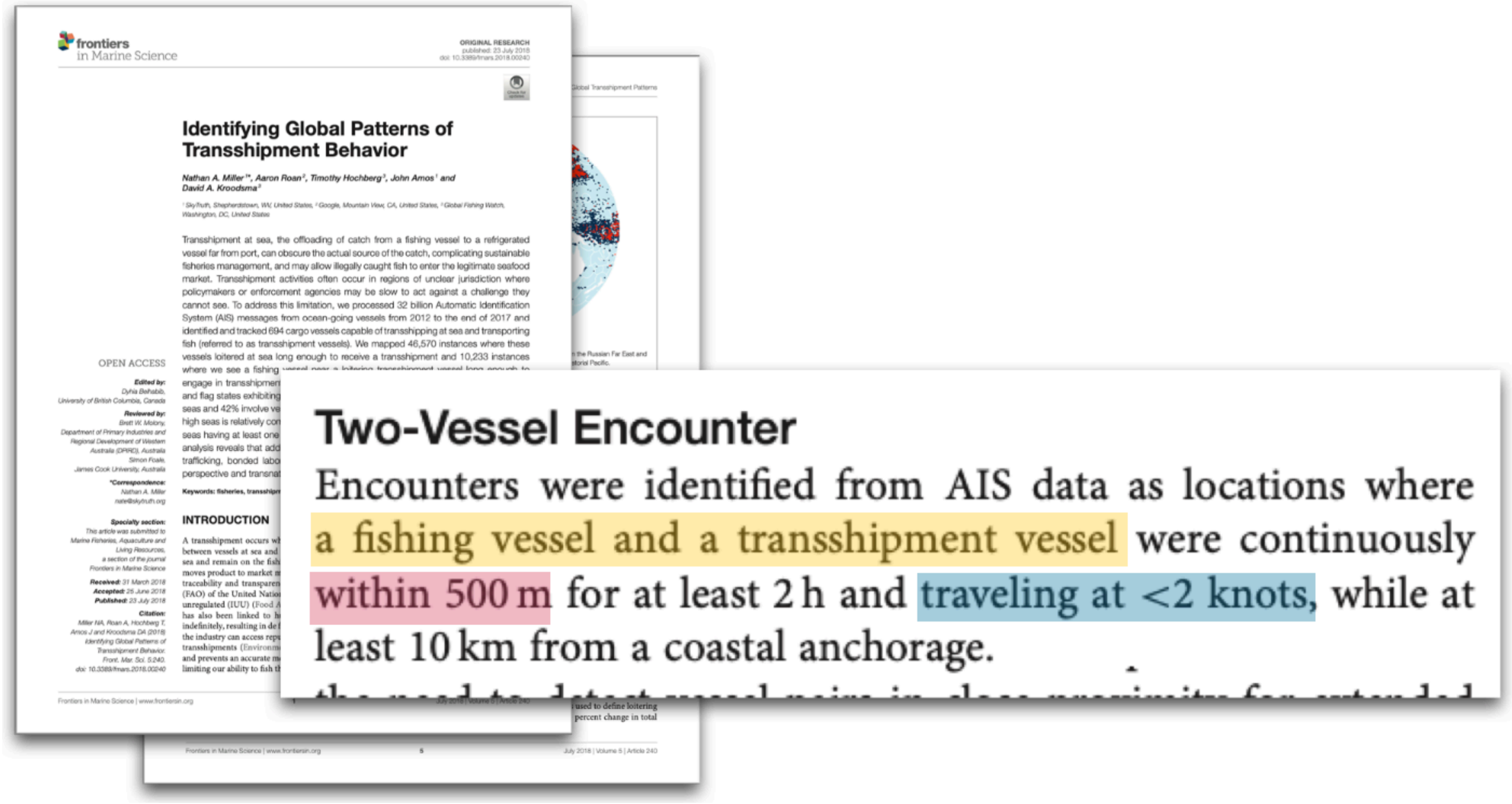

```
→  data git:(master) head reefer2.csv
shipname,mmsi
FRIO FORWIN,210631000
AQUAMARINE,214182223
OCEAN STAR 98,214182700
MINIK ARCTICA,219663000
IVALO ARCTICA,219667000
IZAR ARGIA,224434000
V CENTENARIO,224739000
MARINERO2,238187240
SANTINA,256384000
```

One-time CSV data load into Kafka

```
$ ccloud kafka topic create reefers
$ tr -d '\r' < Reefer.csv | \
  awk -F";" ' { print $2 "\x1c" $1 } '| \
  docker run --rm --interactive edenhill/kafkacat:1.6.0 \
  kafkacat -X security.protocol=SASL_SSL -X sasl.mechanisms=PLAIN \
           -X ssl.ca.location=./etc/ssl/cert.pem -X api.version.request=true \
           -b $CCLOUD_BROKER:9092 \
           -X sasl.username="$CCLOUD_API_KEY" \
           -X sasl.password="$CCLOUD_API_SECRET" \
           -t reefers -K$'\x1c' -P

$ ccloud kafka topic create transshipment-vessels
$ tr -d '\r' < transshipment-vessels-v20170717.csv | \
  sed  "s/,/$(printf '\x1c')/" | \
  docker run --rm --interactive edenhill/kafkacat:1.6.0 \
  kafkacat -X security.protocol=SASL_SSL -X sasl.mechanisms=PLAIN \
           -X ssl.ca.location=./etc/ssl/cert.pem -X api.version.request=true \
           -b $CCLOUD_BROKER:9092 \
           -X sasl.username="$CCLOUD_API_KEY" \
           -X sasl.password="$CCLOUD_API_SECRET" \
           -t transshipment-vessels -K$'\x1c' -P
```

Pattern matching (part 1)



frontiers in Marine Science

ORIGINAL RESEARCH
published: 23 July 2018
doi: 10.3389/fmars.2018.00240

Identifying Global Patterns of Transshipment Behavior

Nathan A. Miller[1]*, *Aaron Roan*[2], *Timothy Hochberg*[3], *John Amos*[1] and *David A. Kroodsma*[3]

[1] SkyTruth, Shepherdstown, WV, United States, [2] Google, Mountain View, CA, United States, [3] Global Fishing Watch, Washington, DC, United States

Transshipment at sea, the offloading of catch from a fishing vessel to a refrigerated vessel far from port, can obscure the actual source of the catch, complicating sustainable fisheries management, and may allow illegally caught fish to enter the legitimate seafood market. Transshipment activities often occur in regions of unclear jurisdiction where policymakers or enforcement agencies may be slow to act against a challenge they cannot see. To address this limitation, we processed 32 billion Automatic Identification System (AIS) messages from ocean-going vessels from 2012 to the end of 2017 and identified and tracked 694 cargo vessels capable of transshipping at sea and transporting fish (referred to as transshipment vessels). We mapped 46,570 instances where these vessels loitered at sea long enough to receive a transshipment and 10,233 instances where we see a fishing

OPEN ACCESS

Edited by:
Dyhia Belhabib,
University of British Columbia, Canada

Reviewed by:
Brett W. Molony,
Department of Primary Industries and Regional Development of Western Australia (DPIRD), Australia
Simon Foale,
James Cook University, Australia

***Correspondence:**
Nathan A. Miller
nate@skytruth.org

Specialty section:
This article was submitted to Marine Fisheries, Aquaculture and Living Resources, a section of the journal Frontiers in Marine Science

Received: 31 March 2018
Accepted: 25 June 2018
Published: 23 July 2018

Citation:
Miller NA, Roan A, Hochberg T, Amos J and Kroodsma DA (2018) Identifying Global Patterns of Transshipment Behavior. Front. Mar. Sci. 5:240. doi: 10.3389/fmars.2018.00240

engage in transshipmen
and flag states exhibiting
seas and 42% involve ve
high seas is relatively con
seas having at least one
analysis reveals that add
trafficking, bonded labo
perspective and transnat

Keywords: fisheries, transshipm

INTRODUCTION

A transshipment occurs wh
between vessels at sea and
sea and remain on the fish
moves product to market m
traceability and transparen
(FAO) of the United Nation
unregulated (IUU) (Food A
has also been linked to hu
indefinitely, resulting in de f
the industry can access repu
transshipments (Environm
and prevents an accurate m
limiting our ability to fish th

Frontiers in Marine Science | www.frontiersin.org

Global Transshipment Patterns

the Russian Far East and
torial Pacific.

used to define loitering
percent change in total

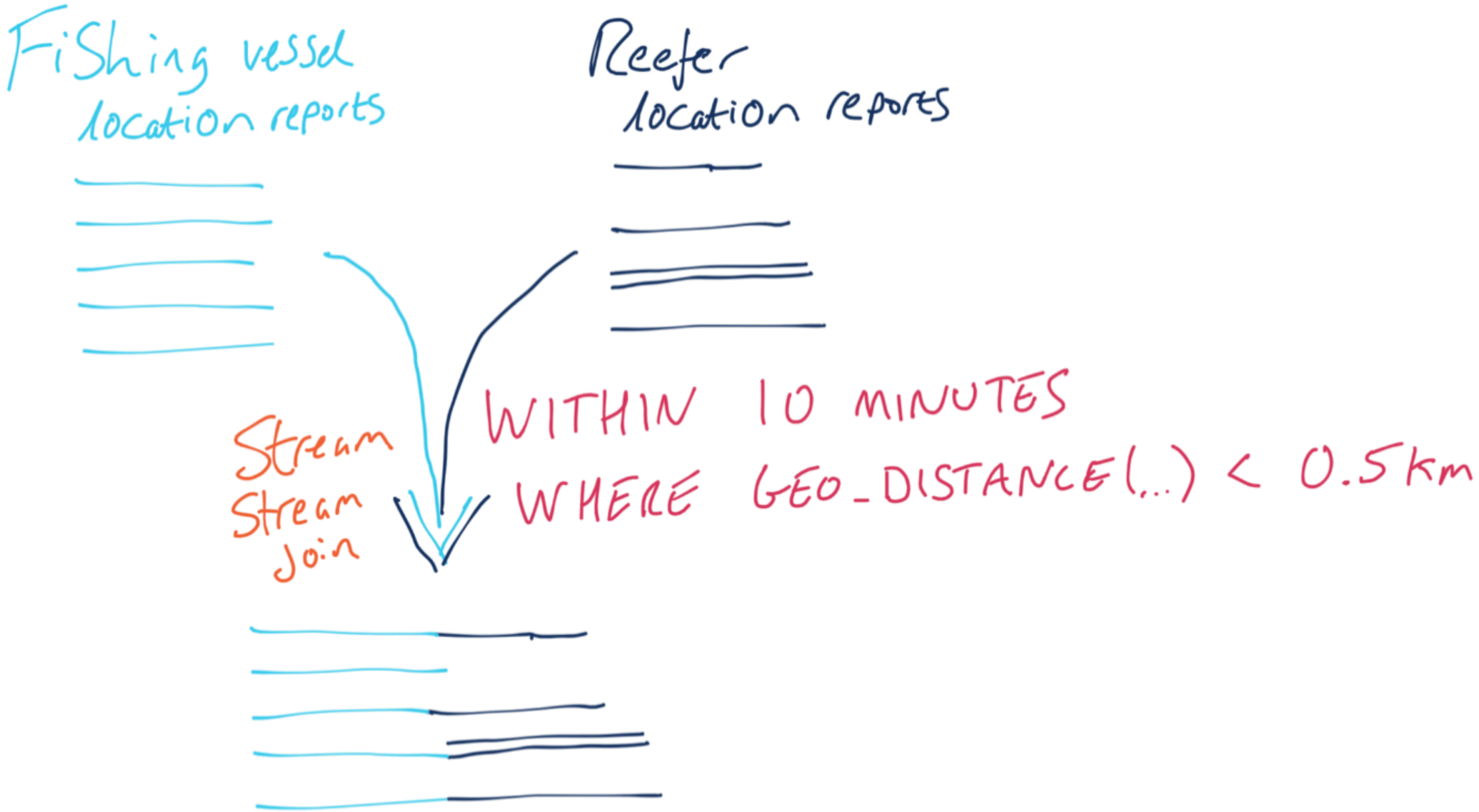
Frontiers in Marine Science | www.frontiersin.org 5 July 2018 | Volume 5 | Article 240

Two-Vessel Encounter

Encounters were identified from AIS data as locations where a fishing vessel and a transshipment vessel were continuously within 500 m for at least 2 h and traveling at <2 knots, while at least 10 km from a coastal anchorage.

Miller NA, Roan A, Hochberg T, Amos J and Kroodsma DA (2018) Identifying Global Patterns of Transshipment Behavior. Front. Mar. Sci. 5:240. doi: 10.3389/fmars.2018.00240

Pattern matching with ksqlDB



Preparing the ship movements stream

```
CREATE STREAM SHIP_STATUS_REPORTS WITH (KAFKA_TOPIC='SHIP_STATUS_REPORTS_V00') AS
SELECT STATUS_REPORT.ROWTIME AS STATUS_TS, STATUS_REPORT.*,  SHIP_INFO.*,
       1 AS DUMMY
FROM  AIS_MSG_TYPE_1_2_3 STATUS_REPORT
      LEFT JOIN SHIP_INFO SHIP_INFO
        ON STATUS_REPORT.MMSI = SHIP_INFO.MMSI;

CREATE STREAM REEFER_MOVEMENTS AS
  SELECT *
    FROM SHIP_STATUS_REPORTS
   WHERE (SHIP_INFO_IS_TRANSSHIPPER=1
       OR SHIP_INFO_IS_REEFER=1);
```

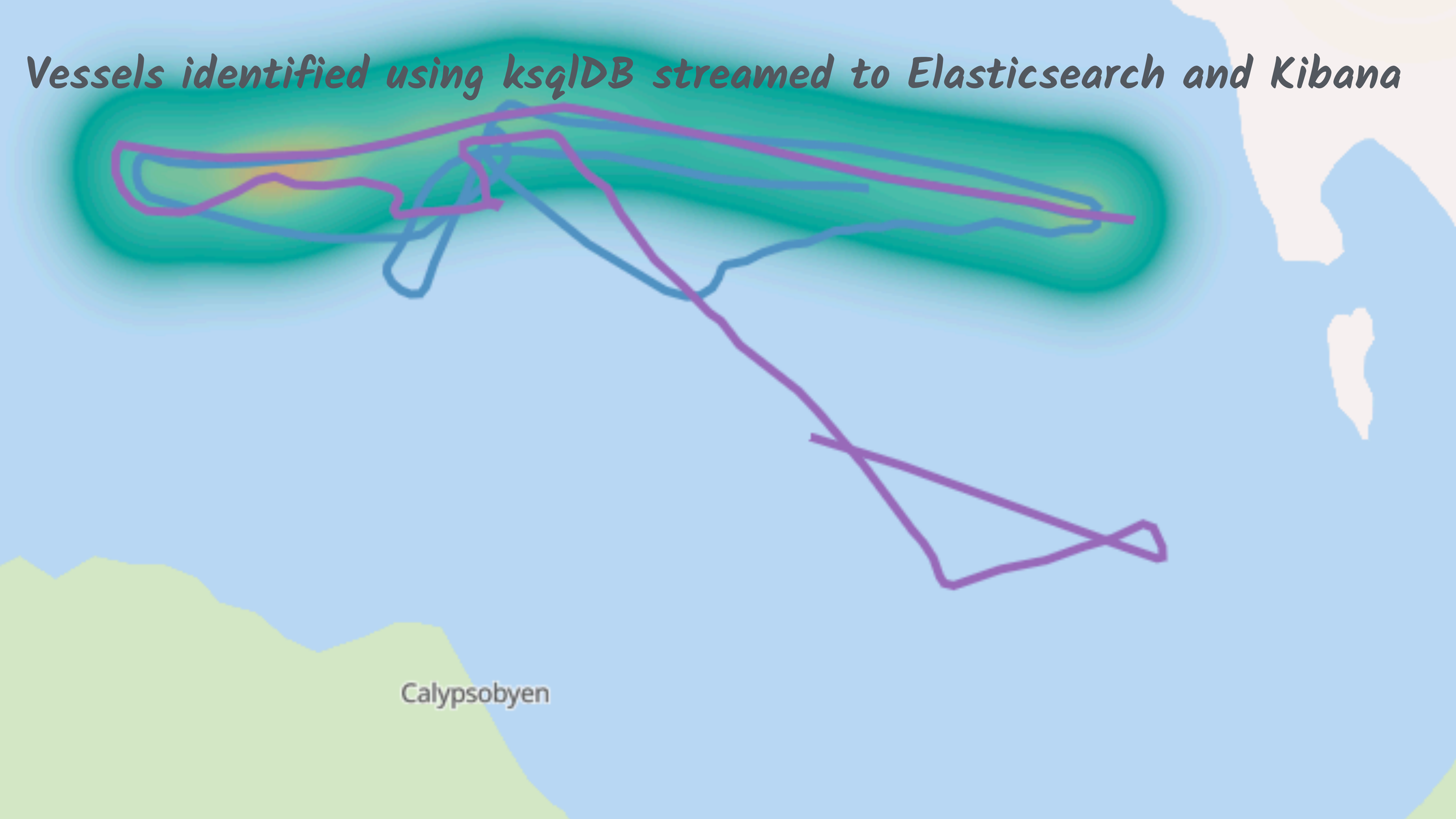
Stream-Stream Join with GEO_DISTANCE

```
CREATE STREAM REEFERS_AND_VESSELS_WITHIN_500M
  WITH (KAFKA_TOPIC='REEFERS_AND_VESSELS_WITHIN_500M_V00') AS
SELECT    V.STATUS_REPORT_MMSI   AS FISHING_VESSEL_MMSI,
          R.STATUS_REPORT_MMSI   AS REEFER_MMSI,
          V.SHIP_LOCATION        AS FISHING_VESSEL_LOCATION,
          R.SHIP_LOCATION        AS REEFER_LOCATION,
          GEO_DISTANCE(V.STATUS_REPORT_LAT, V.STATUS_REPORT_LON, R.STATUS_REPORT_LAT, R.STATUS_REPORT_LON, 'KM') AS
DISTANCE_KM,
          CASE
            WHEN GEO_DISTANCE(V.STATUS_REPORT_LAT, V.STATUS_REPORT_LON, R.STATUS_REPORT_LAT, R.STATUS_REPORT_LON, 'KM') <
0.5
                 AND  (    R.STATUS_REPORT_SPEED <2
                       AND V.STATUS_REPORT_SPEED < 2) THEN 1
                 ELSE 0
             END AS IN_RANGE_AND_SPEED
FROM      SHIP_STATUS_REPORTS V
          INNER JOIN REEFER_MOVEMENTS R
            WITHIN 1 MINUTE
            ON     R.DUMMY = V.DUMMY
WHERE     V.SHIP_INFO_SHIPTYPE_TEXT  = 'Fishing'
AND       V.STATUS_REPORT_MMSI      != R.STATUS_REPORT_MMSI
AND       GEO_DISTANCE(V.STATUS_REPORT_LAT, V.STATUS_REPORT_LON, R.STATUS_REPORT_LAT, R.STATUS_REPORT_LON, 'KM') < 1
PARTITION BY V.STATUS_REPORT_MMSI;
```

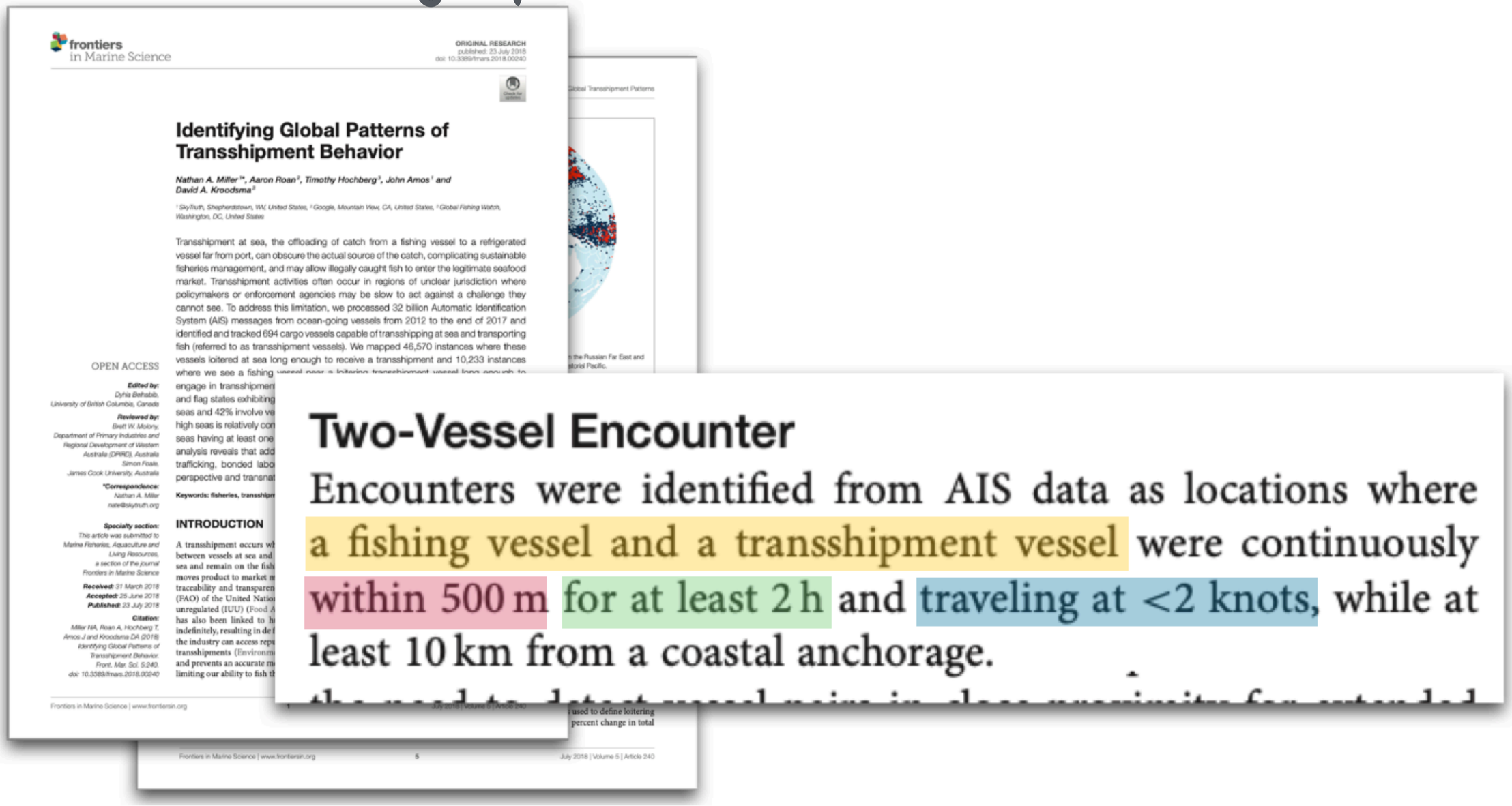
Vessels identified as being near each other

FISHING_VESSEL_MMSI	REEFER_MMSI	FISHING_VESSEL_LOCATION	REEFER_LOCATION	DISTANCE_KM	IN_RANGE_AND_SPEED
258036000	257430000	{lat=62.324023, lon=5.66427}	{lat=62.321373, lon=5.653208}	0.642853083812366	0
273844800	273440860	{lat=69.728918, lon=30.034538}	{lat=69.72819, lon=30.03179}	0.1332701769571591	1
273433220	273440860	{lat=69.72493, lon=30.023542}	{lat=69.72819, lon=30.03179}	0.4820709202116538	1
273433400	273440860	{lat=69.727357, lon=30.03141}	{lat=69.72819, lon=30.03179}	0.09377524348557457	1
257810500	258211000	{lat=62.55565, lon=6.26555}	{lat=62.548723, lon=6.276892}	0.9649975921607864	0

Vessels identified using ksqlDB streamed to Elasticsearch and Kibana



Pattern matching (part 2)



frontiers in Marine Science

ORIGINAL RESEARCH
published: 23 July 2018
doi: 10.3389/fmars.2018.00240

Identifying Global Patterns of Transshipment Behavior

Nathan A. Miller[1]*, Aaron Roan*[2]*, Timothy Hochberg*[3]*, John Amos*[1] *and David A. Kroodsma*[3]

[1] SkyTruth, Shepherdstown, WV, United States, [2] Google, Mountain View, CA, United States, [3] Global Fishing Watch, Washington, DC, United States

Transshipment at sea, the offloading of catch from a fishing vessel to a refrigerated vessel far from port, can obscure the actual source of the catch, complicating sustainable fisheries management, and may allow illegally caught fish to enter the legitimate seafood market. Transshipment activities often occur in regions of unclear jurisdiction where policymakers or enforcement agencies may be slow to act against a challenge they cannot see. To address this limitation, we processed 32 billion Automatic Identification System (AIS) messages from ocean-going vessels from 2012 to the end of 2017 and identified and tracked 694 cargo vessels capable of transshipping at sea and transporting fish (referred to as transshipment vessels). We mapped 46,570 instances where these vessels loitered at sea long enough to receive a transshipment and 10,233 instances where we see a fishing

OPEN ACCESS

Edited by:
Dyhia Belhabib,
University of British Columbia, Canada

Reviewed by:
Brett W. Molony,
Department of Primary Industries and Regional Development of Western Australia (DPIRD), Australia
Simon Foale,
James Cook University, Australia

***Correspondence:**
Nathan A. Miller
nate@skytruth.org

Specialty section:
This article was submitted to Marine Fisheries, Aquaculture and Living Resources, a section of the journal Frontiers in Marine Science

Received: *31 March 2018*
Accepted: *25 June 2018*
Published: *23 July 2018*

Citation:
Miller NA, Roan A, Hochberg T, Amos J and Kroodsma DA (2018) Identifying Global Patterns of Transshipment Behavior. Front. Mar. Sci. 5:240. doi: 10.3389/fmars.2018.00240

engage in transshipment
and flag states exhibiting
seas and 42% involve ve
high seas is relatively co
seas having at least one
analysis reveals that add
trafficking, bonded labo
perspective and transna

Keywords: fisheries, transshipm

INTRODUCTION

A transshipment occurs wh
between vessels at sea and
sea and remain on the fish
moves product to market m
traceability and transparen
(FAO) of the United Natio
unregulated (IUU) (Food A
has also been linked to h
indefinitely, resulting in de
the industry can access repo
transshipments (Environm
and prevents an accurate m
limiting our ability to fish th

Frontiers in Marine Science | www.frontiersin.org

Global Transshipment Patterns

the Russian Far East and
storial Pacific.

used to define loitering
percent change in total

Frontiers in Marine Science | www.frontiersin.org 5 July 2018 | Volume 5 | Article 240

Two-Vessel Encounter

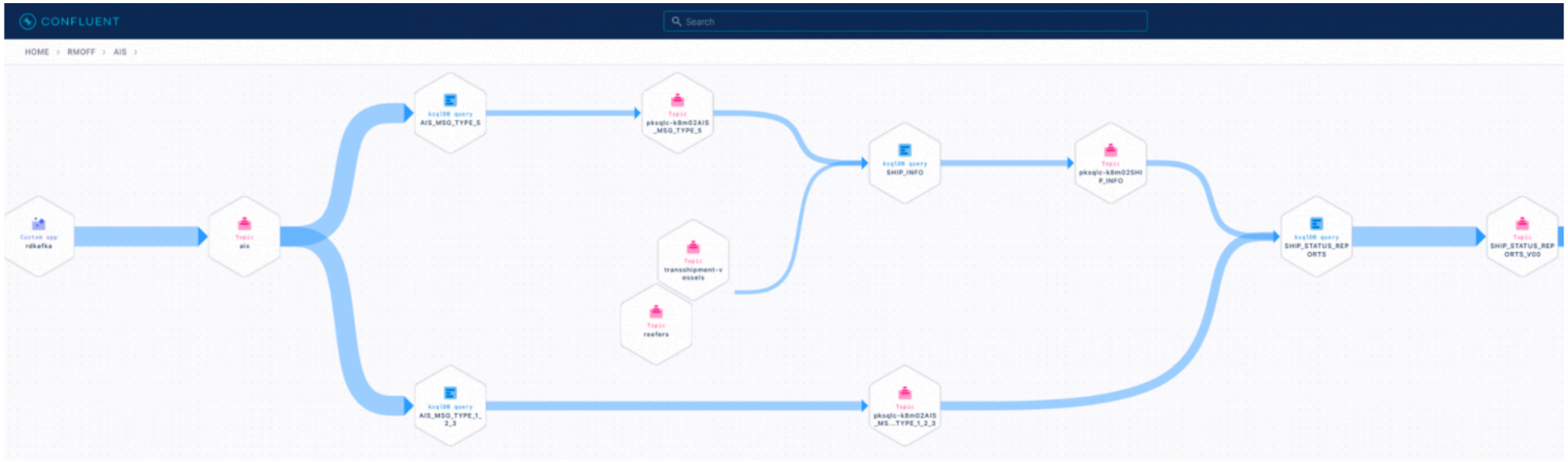
Encounters were identified from AIS data as locations where a fishing vessel and a transshipment vessel were continuously within 500 m for at least 2 h and traveling at <2 knots, while at least 10 km from a coastal anchorage.

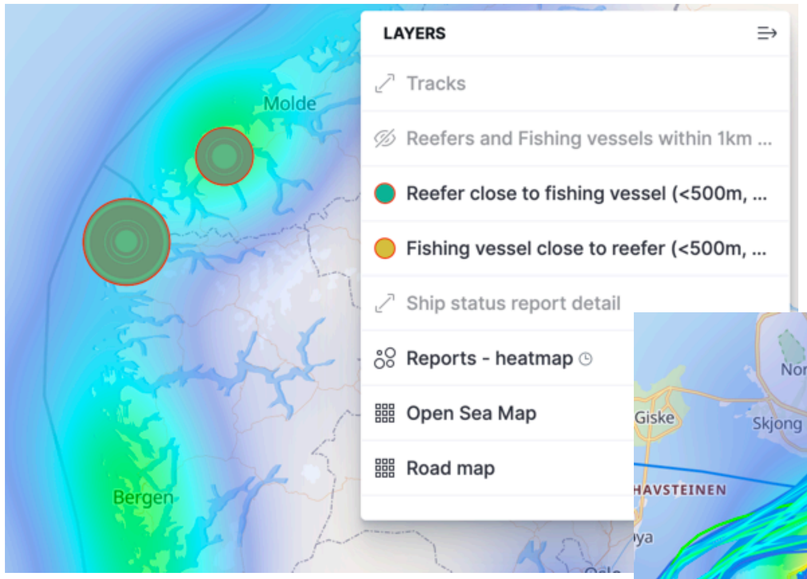
Miller NA, Roan A, Hochberg T, Amos J and Kroodsma DA (2018) Identifying Global Patterns of Transshipment Behavior. Front. Mar. Sci. 5:240. doi: 10.3389/fmars.2018.00240

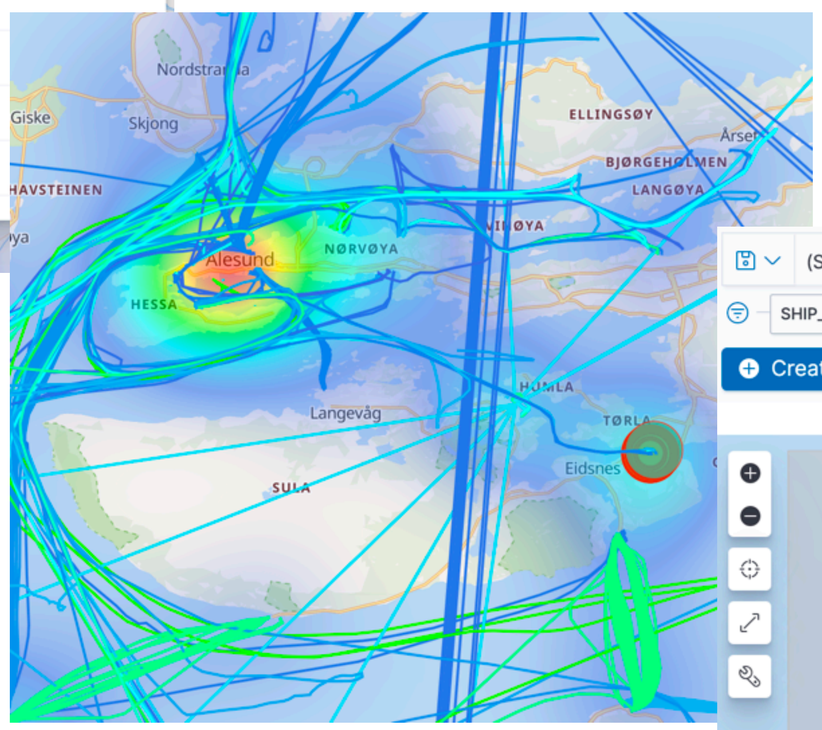
Session windowed aggregations with ksqlDB

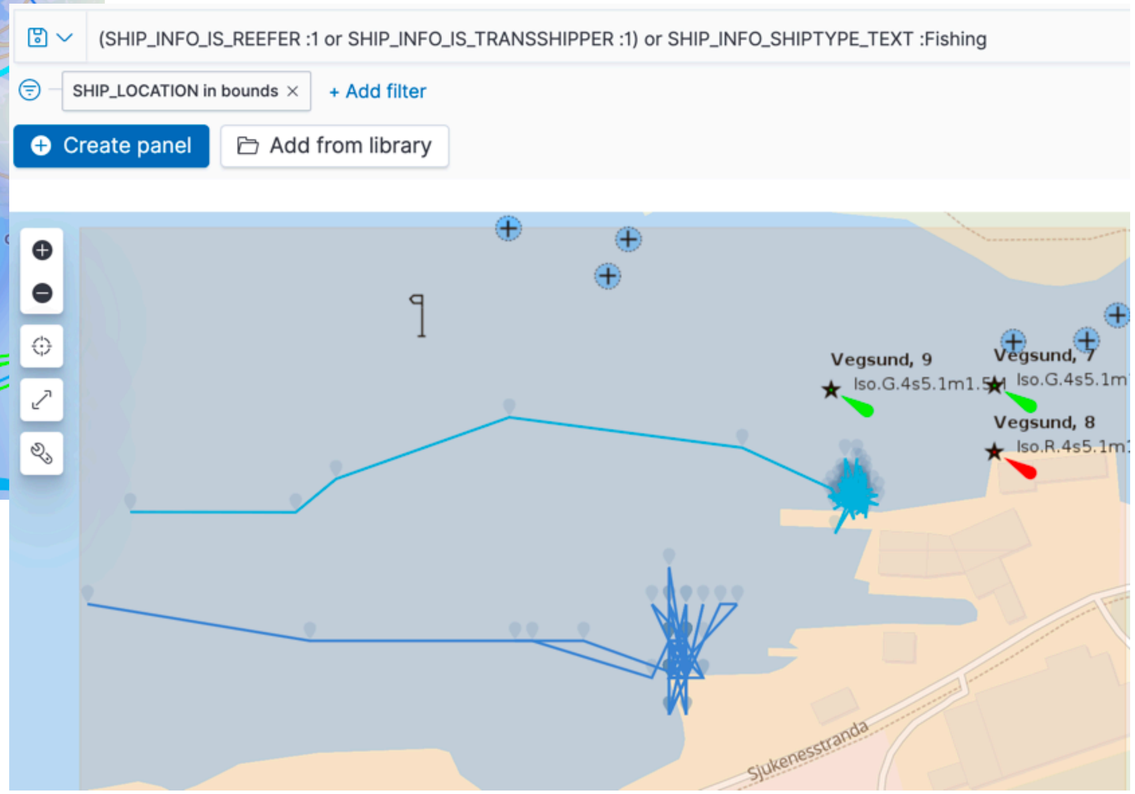
```
SELECT FISHING_VESSEL_MMSI,
       REEFER_MMSI,
       TIMESTAMPTOSTRING(MIN(ROWTIME),'HH:mm:ss','Europe/Oslo') AS FIRST_TIME,
       TIMESTAMPTOSTRING(MAX(ROWTIME),'HH:mm:ss','Europe/Oslo') AS LAST_TIME,
       (MAX(ROWTIME) - MIN(ROWTIME)) / 1000                     AS DIFF_SEC,
       MAX(DISTANCE_KM) AS FURTHEST_DISTANCE,
       COUNT(*) AS EVENT_CT
FROM   REEFERS_AND_VESSELS_WITHIN_1KM WINDOW SESSION (10 MINUTES)
WHERE  IN_RANGE_AND_SPEED=1
GROUP BY FISHING_VESSEL_MMSI, REEFER_MMSI
EMIT CHANGES;
+-------------------+-----------+----------+----------+---------+------------------+---------+
|FISHING_VESSEL_MMSI |REEFER_MMSI |FIRST_TIME |LAST_TIME |DIFF_SEC |FURTHEST_DISTANCE  |EVENT_CT |
+-------------------+-----------+----------+----------+---------+------------------+---------+
|273433220           |273440860   |11:57:14   |12:33:13  |2159     |0.484604604726739 2|13       |
|258036000           |257430000   |13:01:07   |13:04:54  |227      |0.499763431728909 5|4        |
|257888000           |258211000   |11:57:54   |13:33:45  |5751     |0.303924534443280 4|58       |
```

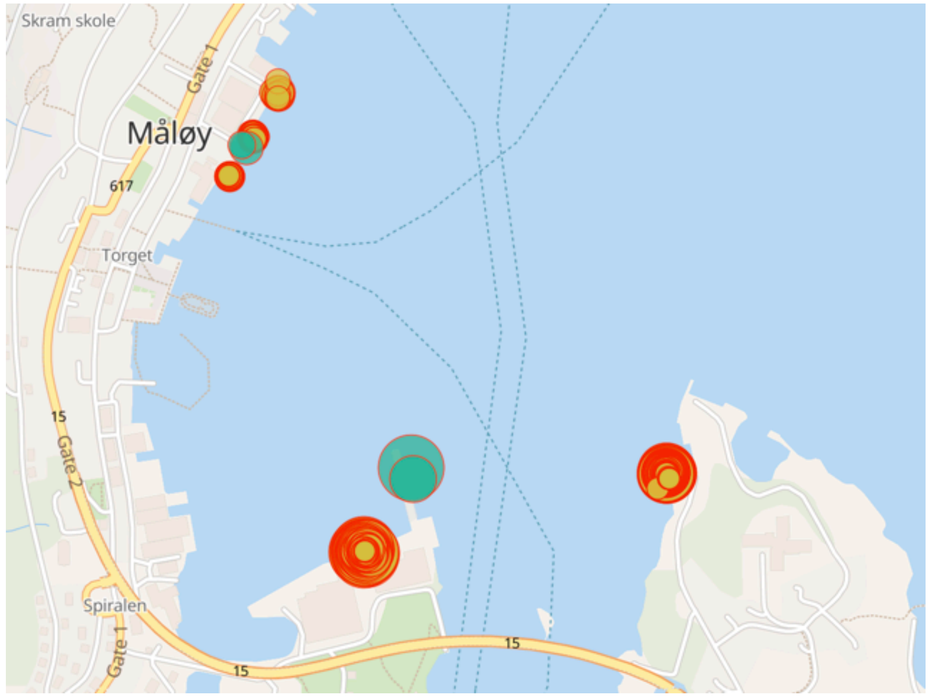
```
CREATE TABLE REEFERS_AND_VESSELS_CLOSE_FOR_GT_2HR
  WITH (KAFKA_TOPIC='REEFERS_AND_VESSELS_CLOSE_FOR_GT_2HR_V00') AS
SELECT   FISHING_VESSEL_MMSI,
         REEFER_MMSI,
         STRUCT("lat":=LATEST_BY_OFFSET(FISHING_VESSEL_LAT),"lon":=LATEST_BY_OFFSET(FISHING_VESSEL_LON)) AS
LATEST_FISHING_VESSEL_LOCATION,
         STRUCT("lat":=LATEST_BY_OFFSET(REEFER_LAT),       "lon":=LATEST_BY_OFFSET(REEFER_LON))         AS LATEST_REEFER_LOCATION,
         MIN(DISTANCE_KM)                                                                               AS CLOSEST_DISTANCE_KM,
         MAX(DISTANCE_KM)                                                                               AS FURTHEST_DISTANCE_KM,
         MIN(FISHING_VESSEL_TS)                                                                         AS FIRST_TS,
         MAX(FISHING_VESSEL_TS)                                                                         AS LAST_TS,
         (MAX(FISHING_VESSEL_TS) - MIN(FISHING_VESSEL_TS)) / 1000                                       AS DIFF_SEC
FROM     REEFERS_AND_VESSELS_WITHIN_1KM WINDOW SESSION (10 MINUTES)
WHERE    IN_RANGE_AND_SPEED = 1
GROUP BY FISHING_VESSEL_MMSI,
         REEFER_MMSI
HAVING   (MAX(ROWTIME) - MIN(ROWTIME)) / 1000 > 7200;
```

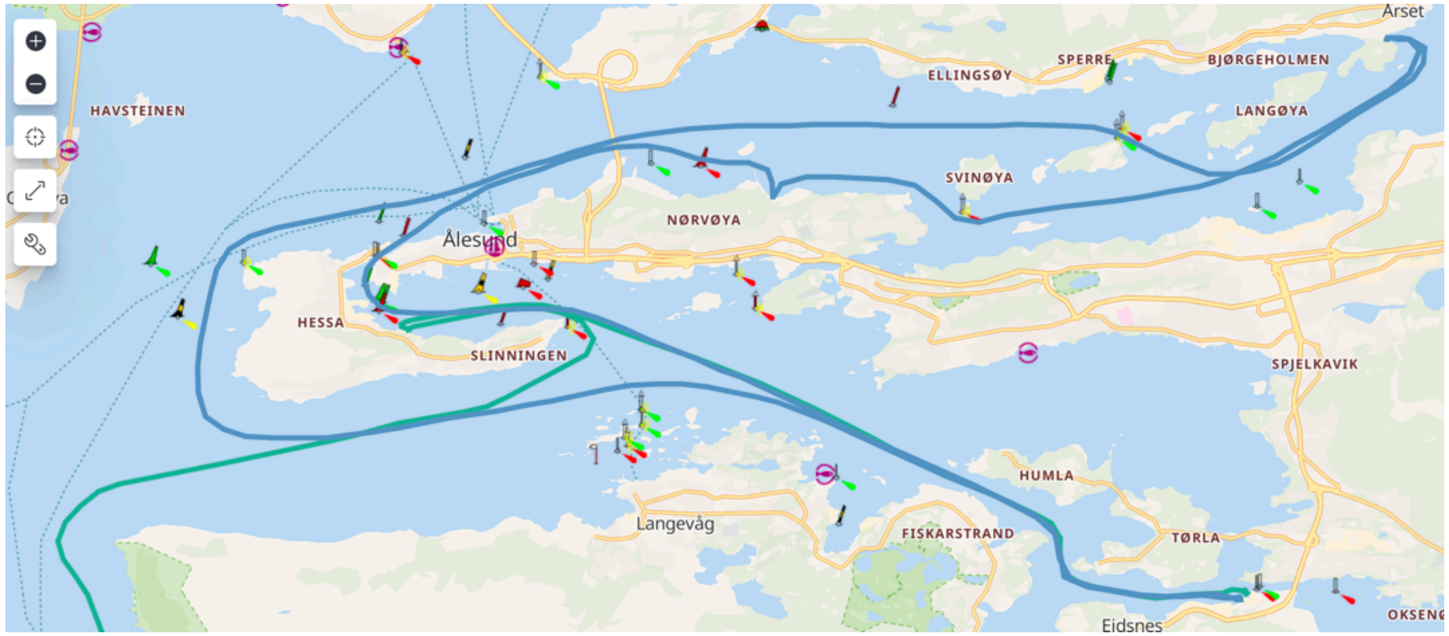



LAYERS
Tracks
Reefers and Fishing vessels within 1km ...
Reefer close to fishing vessel (<500m, ...
Fishing vessel close to reefer (<500m, ...
Ship status report detail
Reports - heatmap
Open Sea Map
Road map
Molde
Bergen

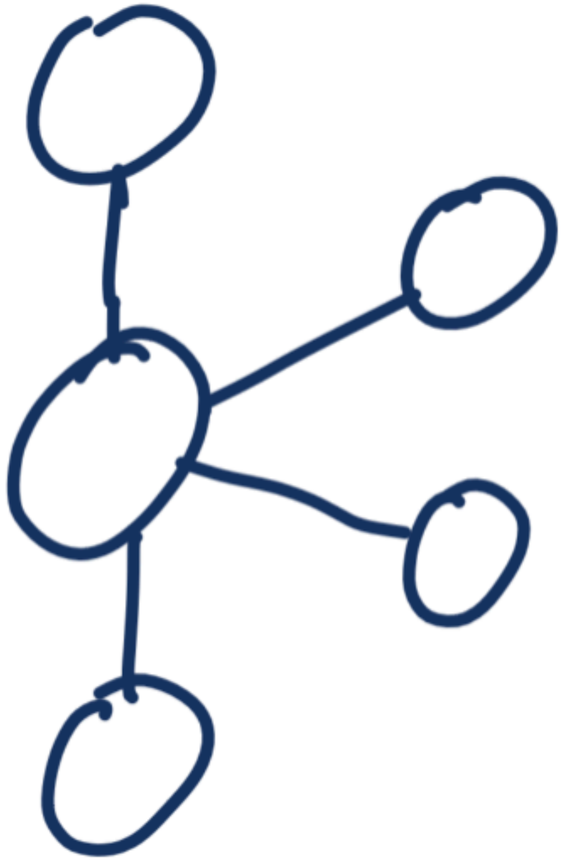

Nordstrandia
Giske
Skjong
ELLINGSØY
Årset
BJØRGEHOLMEN
LANGØYA
HAVSTEINEN
NØRVØYA
Ålesund
HESSA
HUMLA
TØRLA
Langevåg
Eidsnes
SULA

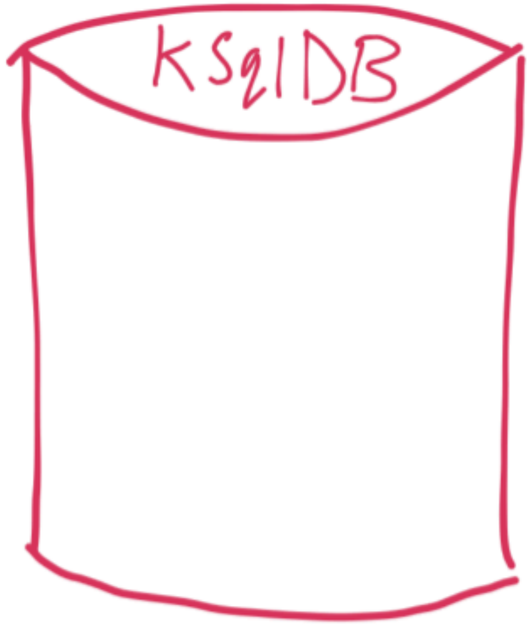

(SHIP_INFO_IS_REEFER :1 or SHIP_INFO_IS_TRANSSHIPPER :1) or SHIP_INFO_SHIPTYPE_TEXT :Fishing
SHIP_LOCATION in bounds
+ Add filter
Create panel
Add from library
Vegsund, 9
Iso.G.4s5.1m1.5M
Vegsund, 7
Iso.G.4s5.1m
Vegsund, 8
Iso.R.4s5.1m
Sjukenesstranda

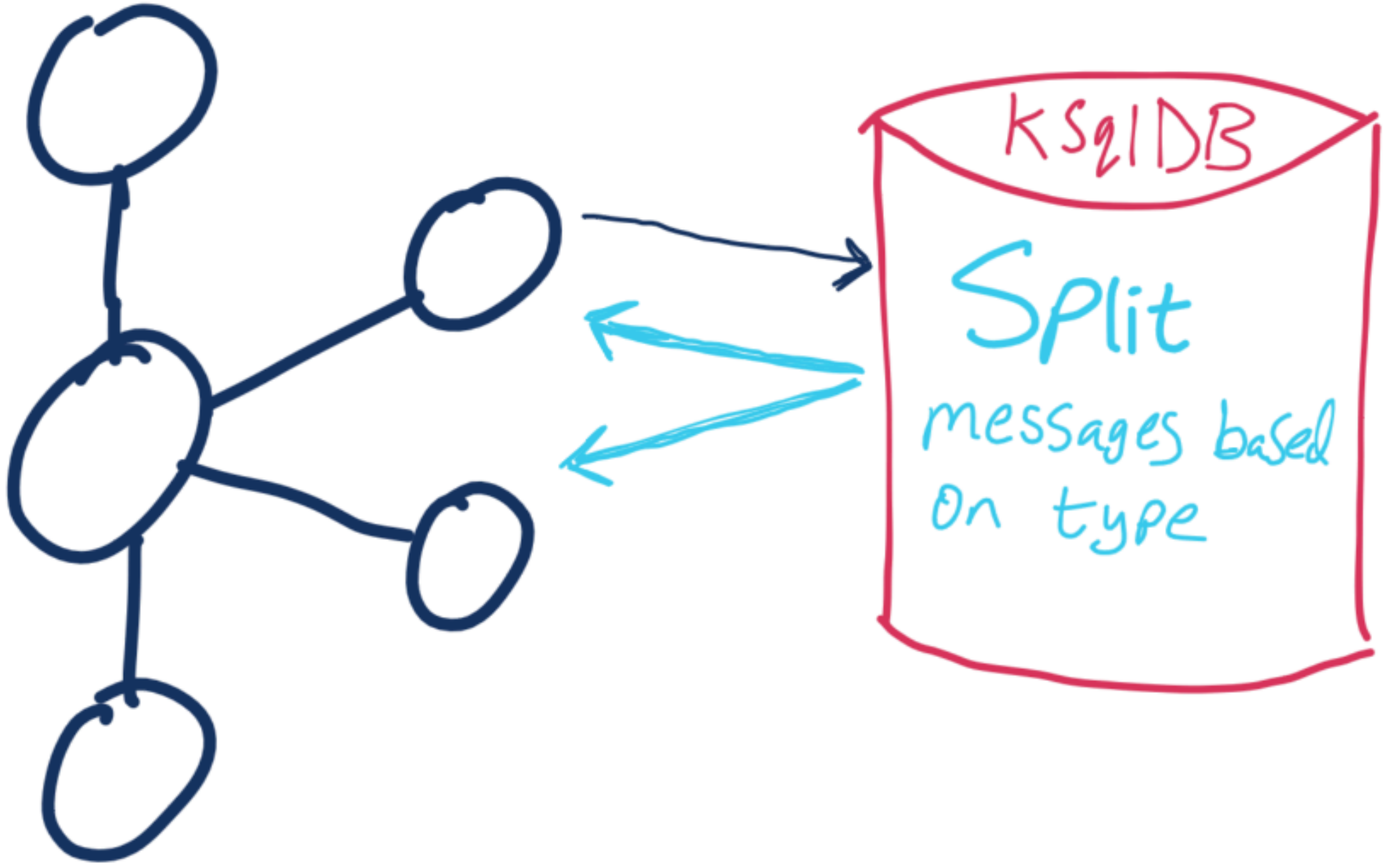

Skram skole
Gate 1
Måløy
617
Torget
15
Gate 2
Spiralen
Gate 1
15
15

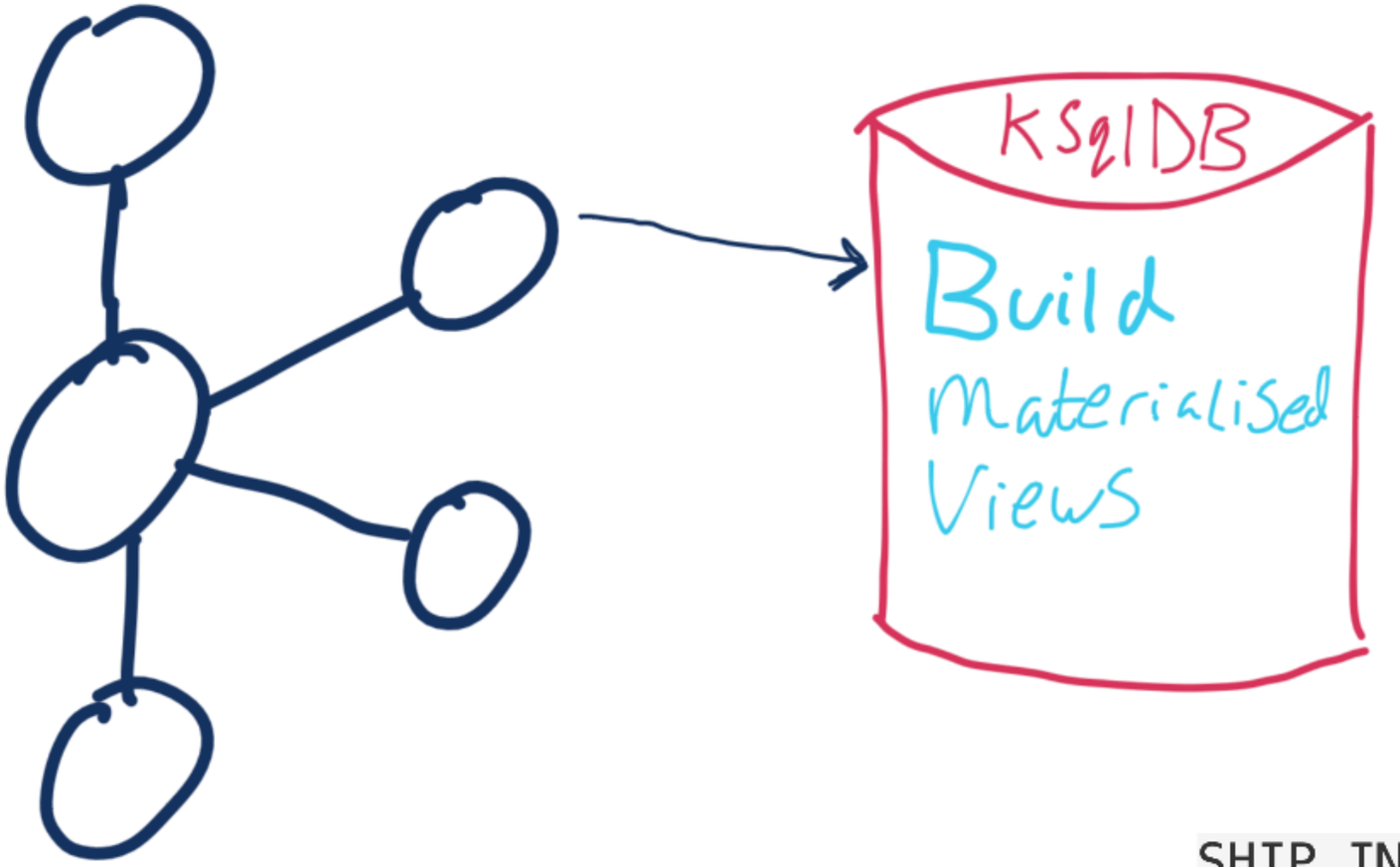

HAVSTEINEN
ELLINGSØY
SPERRE
BJØRGEHOLMEN
LANGØYA
SVINØYA
NØRVØYA
Ålesund
HESSA
SLINNINGEN
SPJELKAVIK
HUMLA
Langevåg
FISKARSTRAND
TØRLA
Eidsnes

Recap

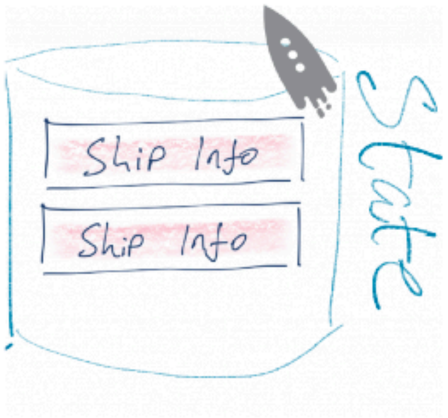


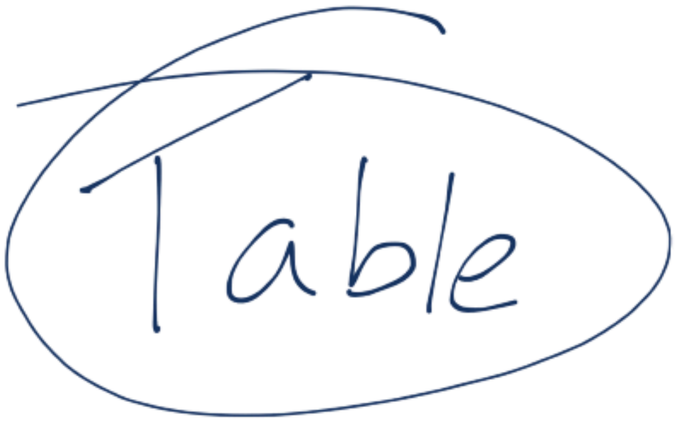

ksqlDB


ksqlDB
Split
messages based
on type

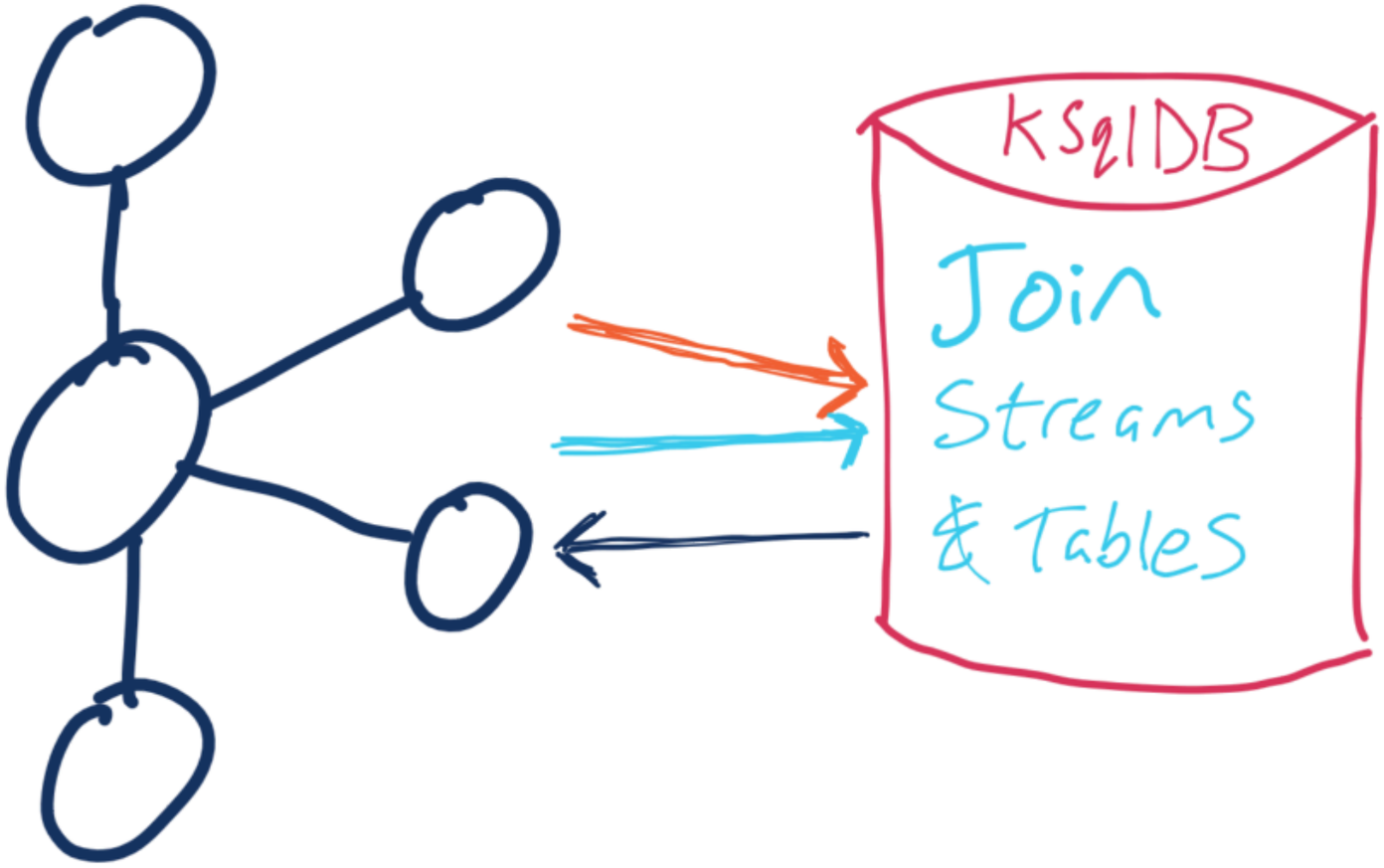

ksqlDB
Build
Materialised
Views

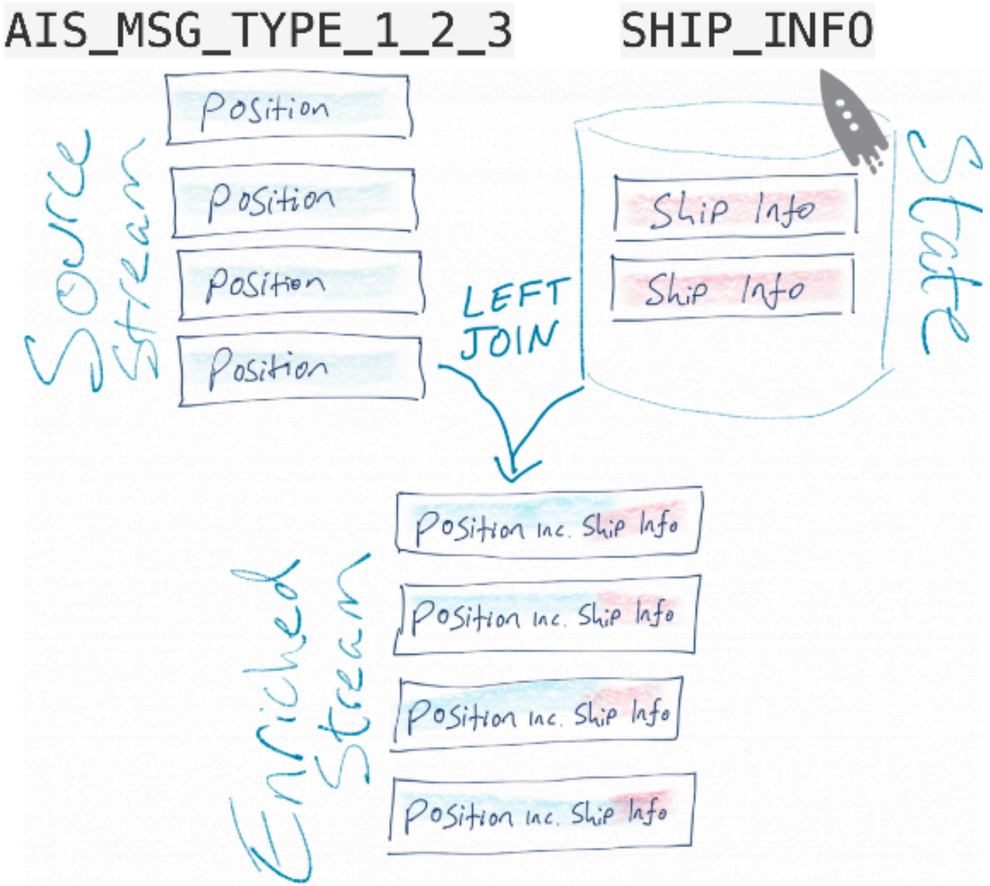
SHIP_INFO

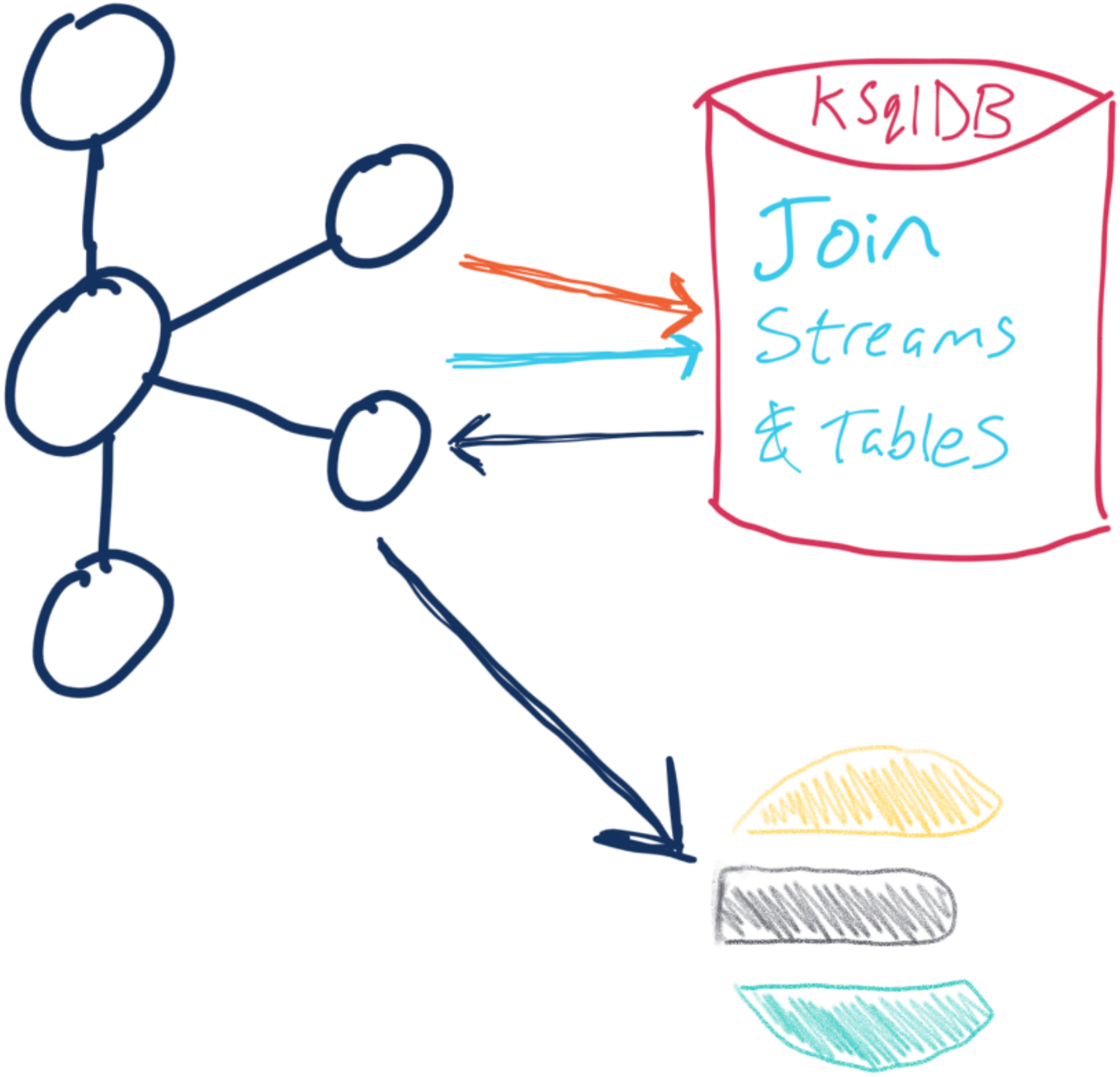

Ship Info
Ship Info
State

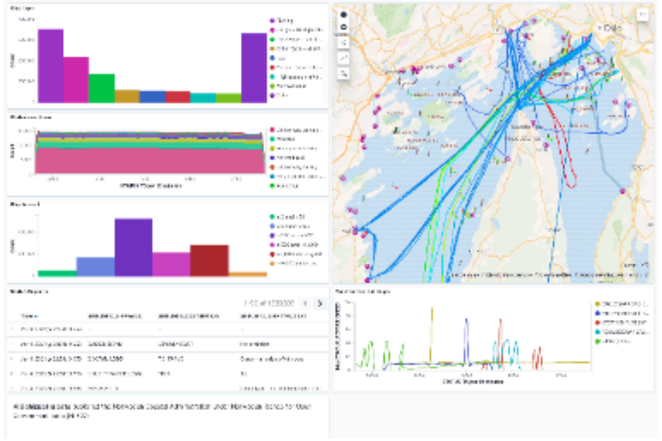


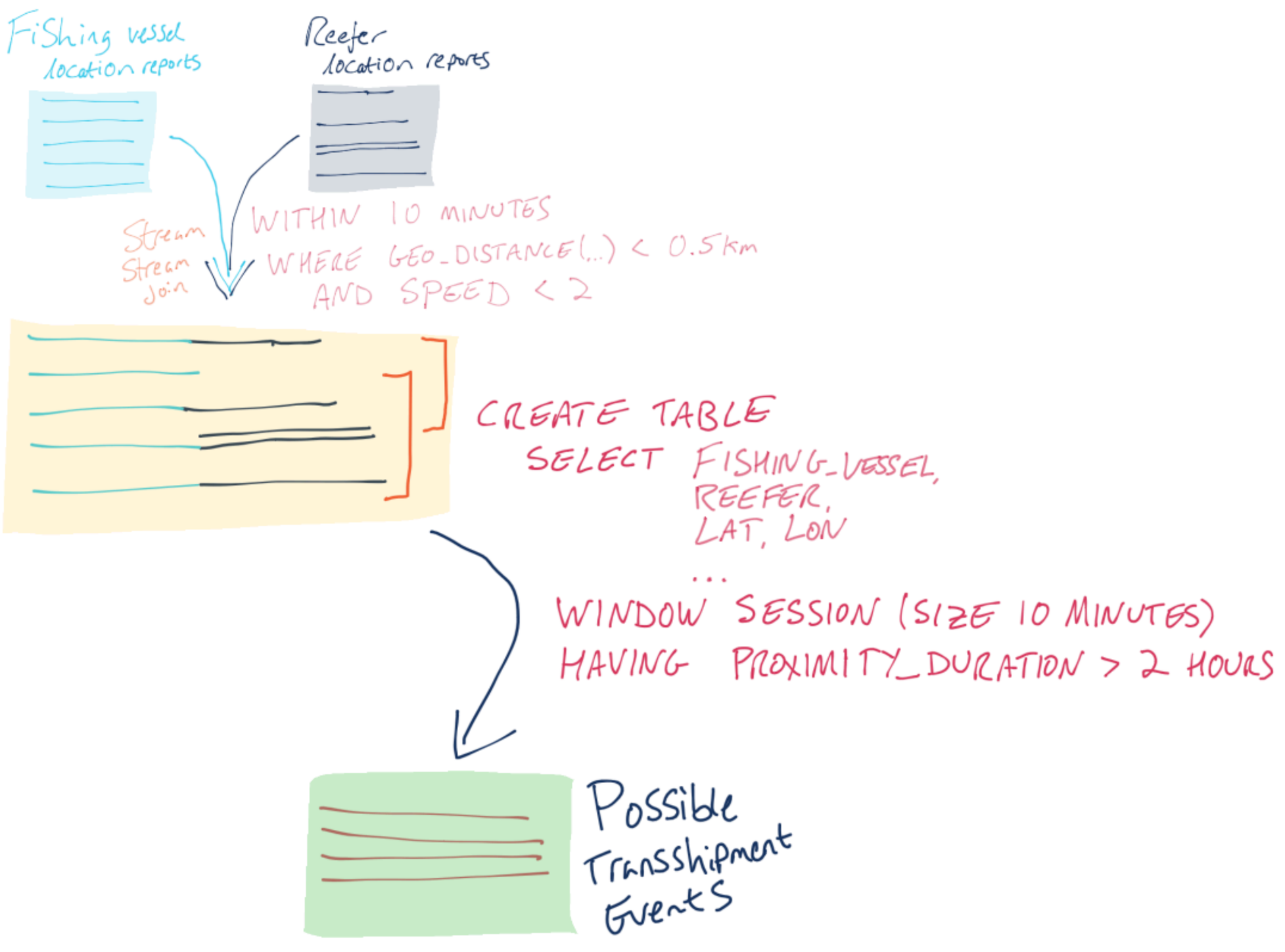
Table

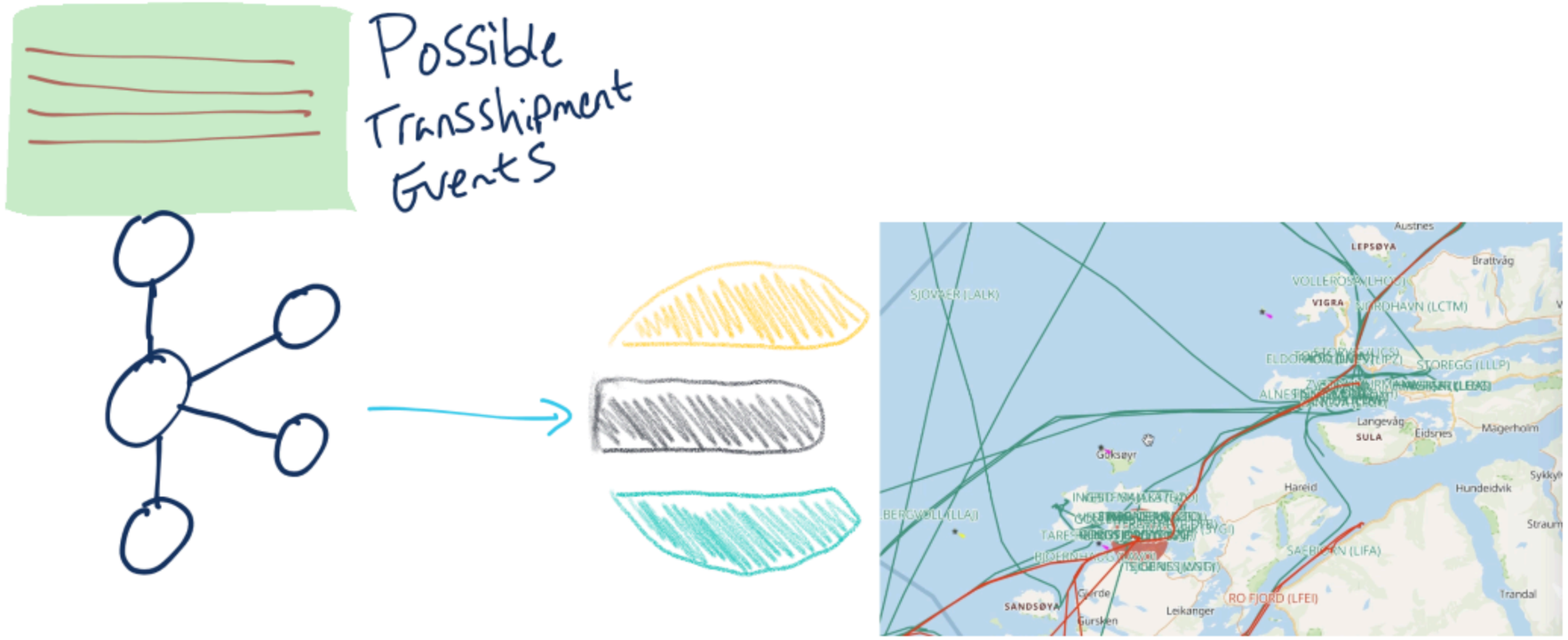

ksqlDB
Join
Streams
& Tables


AIS_MSG_TYPE_1_2_3
SHIP_INFO
Source Stream
Position
Position
Position
Position
LEFT JOIN
Ship Info
Ship Info
State
Enriched Stream
Position inc. Ship Info
Position inc. Ship Info
Position inc. Ship Info
Position inc. Ship Info

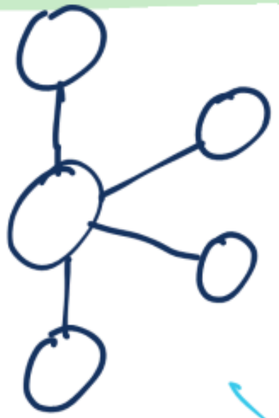

ksqlDB
Join
Streams
& Tables

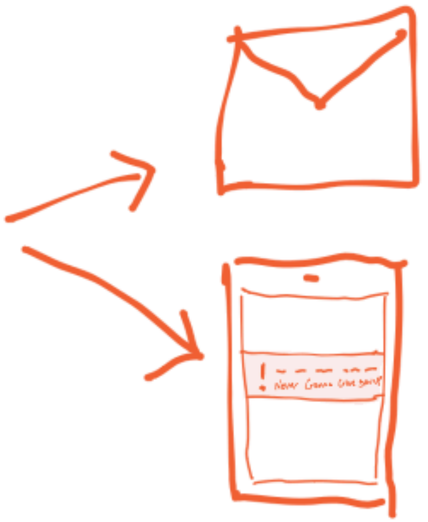



Fishing vessel location reports
Reefer location reports
Stream Stream Join
WITHIN 10 MINUTES
WHERE GEO_DISTANCE(...) < 0.5km
AND SPEED < 2
CREATE TABLE
SELECT FISHING_VESSEL,
REEFER,
LAT, LON
...
WINDOW SESSION (SIZE 10 MINUTES)
HAVING PROXIMITY_DURATION > 2 HOURS
Possible Transshipment Events


Possible
Transshipment
Events
LEPSØYA
Brattvåg
VIGRA
STORDHAVN (LCTM)
STOREGG (LLLP)
Langevåg
SULA
Eidsnes
Magerholm
Haneid
Hundeidvik
SAFEBORN (LIFA)
RO FJORD (LFEI)
SANDSØYA
Leikanger
Trandal

Possible Transshipment Events



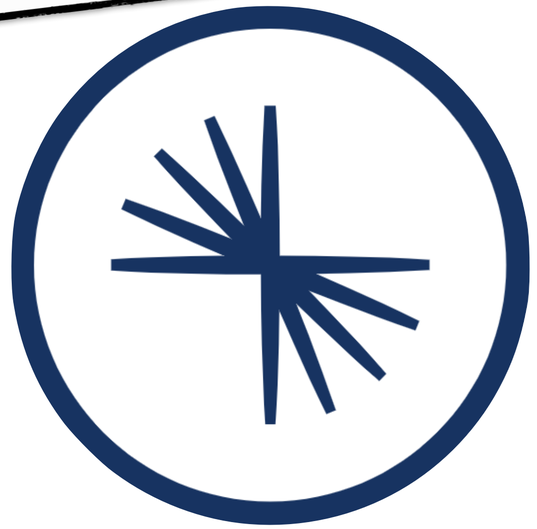


Deep-dive articles

- What is Apache Kafka?
- Event Streaming vs. Related Trends
- KRaft: Kafka without ZooKeeper
- Transactions & Guarantees in Kafka
- Processing & Storage Fundamentals
- Kafka Performance
- Cloud-native Kafka
- Streaming Database Systems
- Testing Apache Kafka
- Explore Kafka's Internals

Over 10 hours of free courses

- Apache Kafka 101
- Kafka Connect 101
- Kafka Streams 101
- ksqlDB 101
- Inside ksqlDB
- Spring Framework and Kafka
- Building Data Pipelines with Kafka
- Event Sourcing with Kafka
- Data Mesh 101



CONFLUENT

Developer

Plus: Hands-on Quick Starts and Client Language Guides + Event Streaming Patterns + More

developer.confluent.io

#EOF

@rmoff

http://youtube.com/rmoff